Characterisation of putative metal transport proteins in the nickel hyperaccumulator *Senecio coronatus*: investigating candidate genes for nickel tolerance and accumulation.

Ross Martin Cowlin

Thesis presented for the degree of Master of Science in the Department of Molecular and Cell Biology at the University of Cape Town

Supervisor: Dr Robert Ingle

April 2017

PLAGIARISM DECLARATION

1. I understand the meaning and implications of plagiarism and declare that all the work in this thesis, other than that properly acknowledged, is my own.

2. I have utilised the Oxford University Press (HUMSOC) conventions for citation and referencing. Each contribution to, and quotation in, this thesis from the work of other people has been acknowledged, and has been appropriately cited and referenced.

3. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his/her own.

Signed by candidate

Signature removed

Ross Martin Cowlin
Signed 31 March 2017

ACKNOWLEDGEMENTS

There are so many people to whom I am deeply indebted for all the support and assistance I have received during the course of this masters. Firstly, thanks to my supervisor Rob Ingle for taking a chance on a humble zoologist. I am very grateful to have been given the opportunity. To the other members of Lab 430: thanks for everything and all the best for the future. Nina, best of luck for the next two years – I know you'll do brilliantly. I am extremely grateful to Lara Donaldson, who has played a huge role in this dissertation. I cannot tell you how much it was appreciated. Thank you too to both UCT and the NRF.

I would like to thank the many academics and staff members in the Department with whom I have interacted – many of whom have assisted me in some way or another. Shane Murray, Laura Roden, Suhail Rafudee, Colleen O'Ryan, Ann Meyers and Shakiera Sattar. I am also grateful to all the incredibly friendly, approachable and hard-working scientific officers, departmental assistants and cleaning staff. Also the administrative staff. A special thanks to Michael Gcwabe, I am really going to miss our little chats in the MCB hallways. My next "zamalek" will be raised in a toast to you.

Neil Bredekamp and Derick September, you guys are awesome. Thank you both for your patience and assistance with the many, many strange requests that I have sent your way. Not sure how I'm going to manage without you when I leave the department. Blommie Filmer and Peter Louw, thank you both so much. I don't think I once left the stores without feeling more cheerful than I had been when I walked in. Oom Pete, I still can't get my head around your commitment to this Department and UCT. All the very best for your approaching retirement. Faezah Davids and Madhu Chauhan, I can't thank you two enough. I honestly don't think this department could continue to function for more than a few hours without you. Madhu, I have really enjoyed hanging out recently. I will miss you both. Thank you for everything. Keren Cooper, you've been a fantastic friend and colleague and I have really enjoyed getting to know you.

There are many people I need to thank outside MCB. Thank you to Jonathan Colville and Annelie Melin, hopefully we will see more of each other as I close this chapter. Ren Larison, thank you for your words of encouragement at various moments of doubt. Janine Cowlin, thank you for your patience and support over the years. To the "Brohemes", thanks for all the good times and many laughs. Seano, thanks so much for that Photoshop tutorial. You had better watch out - I might start poaching your clients. Deon, thank you for the many, many, many beers and the many long and involved discussions about MMA and music. Not much I need beyond those three things really. Nicole, you helped me out of a really tight spot. I am so grateful. So is Mirko. Phil, same to you boet. I really appreciate you bailing me out like you did. Sarah too. I am very grateful to have friends like you in my life.

Blaine, as much as I hate to say this, I think you are one of the few reasons I have managed to maintain my sanity. You just keep dragging me back. Infuriating. But thank you so much for your friendship. I am excited about the next chapter of your life, mental though it is no doubt going to be. Speaking of both sanity and chapters, I cannot fail to acknowledge Wodehouse. Definitely another big factor in keeping me more-or-less sane. Pritchard, you should try harder with him.

Walt, thank you so much for everything. The last few years have been, well, tough, but you've always been there for me. The German Toilet Incident. The Jackson Pollock Bathroom Incident. Etc. I am especially grateful for your support during the last few months of chaos – and welcoming me into your home during trying times of your own. I'm not sure I'll ever be able to repay you (I mean metaphorically – paying you back for all the pizza you've bought me over the years is a ship that has very definitely sailed), but I will do what I can. I really couldn't ask for a better friend.

Mom, thank you for your incredible support over the last few years. I know I can be, well, curmudgeonly, but I hope you know how much I appreciate your support. I probably haven't said it enough, but I am really proud of your achievements. Dad, I really can't thank you enough for everything you do and have done for me. I honestly don't know what I would do without you. The last few years have been rough, but you've never once let me down. Thank you so much for that. I do, however, have a sneaking suspicion that you still think my studies have consisted entirely of learning animal names, but one of these days I'm going to explain it all to you over a beer or three.

I cannot finish writing this without thanking Mirko. I suppose it is part of the tragedy inherent to a man's relationship with his dog that you will never understand how grateful I am to have you in my life. Here's to many more years together, my little fat monkey boy.

Lastly, thank you to Oom Charlie who started it all. I still remember the first time I heard about the book that changed everything. I cannot imagine the world we would be living in today if you hadn't done what you did. Thank you.

Thus, from the war of nature, from famine and death, the most exalted object which we are capable of conceiving, namely, the production of the higher animals, directly follows. There is grandeur in this view of life, with its several powers, having been originally breathed into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved.

~ Charles Darwin, On the Origin of Species

ABSTRACT

The accumulation of exceptionally high concentrations of heavy metals in plant tissues is an extreme phenotypic trait that has evolved independently in multiple plant taxa. The majority of research undertaken in this area has been performed on zinc/cadmium hyperaccumulators and comparatively little is known about the molecular mechanisms behind nickel accumulation. This is despite the fact that nickel hyperaccumulators constitute more than 75% of all known hyperaccumulator species. One such species is *Senecio coronatus* (Asteraceae), which is a useful model to study nickel hyperaccumulation - as both hyperaccumulator and non-accumulator populations have been identified on nickel-rich serpentine soils in South Africa. The nickel-transporting abilities of three proteins (ScMATE, ScVIT and ScCOP), previously shown to be constitutively over-expressed in shoot tissues of hyperaccumulating populations of *S. coronatus*, were investigated in order to determine if they play a role in nickel hyperaccumulation. The RNA-Seq derived nucleotide sequences of these genes were confirmed by reverse transcriptase PCR, and computational analysis suggested that the proteins encoded by these genes display identical topology to their homologues in *Arabidopsis thaliana*. Heterologous expression of these proteins in a metal-sensitive yeast strain was performed to determine whether they are capable of transporting nickel. Although a minor reduction in nickel sensitivity was observed in yeast expressing ScMATE, and a minor increase in ScCOP-expressing yeast, no marked changes in sensitivity to nickel were observed. C-terminal EYFP-tagged MATE and VIT fusion proteins were transiently expressed in live onion cells to determine the subcellular localization of these proteins *in planta*. Fluorescence microscopy indicated that MATE localises to the nucleus and VIT to the tonoplast or plasma membrane.

ABBREVIATIONS

°C	-	degrees Celsius
At-	-	*Arabidopsis thaliana*
bp	-	base pairs (nucleotide)
cDNA	-	complementary deoxyribonucleic acid
COPT/COP	-	copper transporter protein family
CSM-uracil	-	yeast synthetic drop-out medium minus uracil
DNA	-	deoxyribonucleic acid
g	-	gram
GOI	-	gene of interest
kb	-	one thousand base pairs ("kilobase")
kDa	-	kilodalton
L	-	litre
LB	-	lysogeny broth / Luria-Bertani medium
m-	-	milli-
-m	-	-metre
M	-	molar
MATE	-	multidrug and toxin extrusion protein family
ORF	-	open reading frame
RNA	-	ribonucleic acid
RPM	-	revolutions per minute
Sc-	-	*Senecio coronatus*
VIT	-	vacuolar iron transporter protein family
YFP/EYFP	-	yellow fluorescent protein / enhanced YFP
YNB	-	yeast nitrogen base (without amino acids)
YPD	-	yeast extract, peptone, dextrose
µ-	-	micro-
v/v	-	volume per volume
w/v	-	weight per volume

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I INTRODUCTION

1. Overview

The accumulation of exceptionally high concentrations of otherwise toxic metals in plant tissues, termed hyperaccumulation (Brooks et al. 1977), is an extreme phenotypic trait that has evolved independently in multiple plant taxa (Antonovics et al. 1971; Gartside & McNeilly 1974; Krämer 2010; Rascio & Navari-Izzo 2011). Frequently associated with serpentine soils, the concentration of metals in these plants may be up to four orders of magnitude higher than in related non-accumulator taxa (Baker & Brooks 1989; Krämer 2010; Reeves & Baker 2000). Although metal tolerant plants were described as long ago as the sixteenth century (Caesalpino 1583; Thalius 1588), it was not until the middle of the twentieth century that certain species were shown to storing these metals in their tissues (Minguzzi & Vergano 1948). Both the molecular mechanisms and the selective advantage/s of the phenomenon are poorly understood, despite their profound implications for our understanding of metal homeostasis in plants (Krämer 2010; Rascio & Navari-Izzo 2011).

2. Hyperaccumulation of heavy metals

Of the seventeen elements considered essential for plant growth and reproduction (DalCorso et al. 2014; Mengel et al. 2001; Williams & Salt 2009), nine[1] are considered macronutrients (i.e. constitute more than 0.1% plant dry mass) whereas the remaining eight[2] are classified as micronutrients or trace elements (Williams & Salt 2009). Plants, being sessile organisms, have had to evolve complex strategies for the acquisition and maintenance of these elements (Shitan et al. 2014) and, conversely, to prevent over-accumulation thereof (DalCorso et al. 2014; Diener et al. 2001; Wink 1997). The majority of the micronutrients are metals or metalloids (to be referred to generically here as "heavy metals") and are essential to many central aspects of plant functioning, including primary and secondary metabolism, gene regulation,

[1] Carbon, hydrogen, oxygen, nitrogen, sulfur, phosphorus, calcium, potassium and magnesium.
[2] Nickel, molybdenum, copper, zinc, manganese, boron, iron and chlorine.

hormone perception, signal transduction and, because of their redox potential, are vital cofactors in a variety of enzymatic reactions (DalCorso et al. 2014). However, supra-optimal levels of the micronutrient heavy metals have profound phytotoxic effects such as the generation of reactive oxygen species (leading to oxidative stress), biomembrane deterioration and chlorosis (DalCorso et al. 2014; Nagajyoti et al. 2010; Yadav 2010). This is also the case for metals and metalloids not known to have a role in plant functioning, such as arsenic, lead and cadmium (DalCorso et al. 2014), but which may nonetheless be accumulated in plant tissues (Williams & Salt 2009). Nickel, in particular, is important for maintaining urease function in plant leaves, although supra-optimal concentrations lead to reduced root growth and functioning (DalCorso et al. 2014).

There are approximately 295 000 described angiosperm species (Christenhusz & Byng 2016), the majority of which utilise mechanisms to prevent unwanted metal ions from entering the aboveground tissues through complexation and storage in root apoplast and cell walls (Rascio & Navari-Izzo 2011). Approximately 500 angiosperm species, representing just 0.2% of the total, take a contrary strategy (Krämer 2010; Rascio & Navari-Izzo 2011). This strategy, termed hyperaccumulation (Brooks et al. 1977), involves the active accumulation of supra-optimal, typically toxic concentrations of heavy metals. Hyperaccumulation comprises three key features: enhanced uptake of soil metals by the roots, rapid and effective metal ion translocation, efficient detoxification and sequestration of metals in leaf tissues (Krämer 2010; Rascio & Navari-Izzo 2011).

The threshold values used in the determination of a species' status as a hyperaccumulator depend on the metal concerned, but vary from as little as 0.01% (dry weight) for cadmium and up to 1% for manganese or zinc (Rascio & Navari-Izzo 2011). Nickel hyperaccumulators are designated as such when capable of storing upwards of 0.1% nickel (dry weight) in the aboveground tissues, although, incredibly, one species has been shown to accumulate 26% nickel (dry weight) in certain tissues (Sagner et al. 1998). Typically, just 0.005% nickel would result in phytotoxicity (Krämer

2010; Rascio & Navari-Izzo 2011). Despite the fact that nickel hyperaccumulators constitute more than 75% of all known hyperaccumulator species, the vast majority of experimental work done has been performed on zinc/cadmium hyperaccumulators *Arabidopsis halleri* (Brassicaceae) and *Noccaea caerulescens* (Brassicaceae) and comparatively little is known about the molecular mechanisms behind nickel transport and accumulation (Rascio & Navari-Izzo 2011).

3. The evolution of hyperaccumulation

Although not directly relevant to the objectives or outcomes of this study, no discussion regarding the phenomenon of hyperaccumulation would be complete without briefly considering the broader evolutionary context thereof. More specifically, given the obvious metabolic and energetic disadvantages of hyperaccumulation, one would tend to hypothesise that some or other selective advantage/s must exist to balance the equation. It is worth bearing in mind that the trait has evolved on multiple occasions across multiple, unrelated plant taxa and is thus unlikely to be merely an anomalous evolutionary dead-end. Five hypotheses have been proposed to explain the evolution of hyperaccumulation (Boyd & Martens 1992): inadvertent uptake, tolerance or disposal of metal, drought resistance, interference (by competition or allelopathy) and defence against herbivores and/or pathogens. More recently, only the latter two hypotheses have received serious attention and currently compete as explanations for this trait (Boyd & Martens 1992; Boyd et al. 2002; Huitson & Macnair 2003; Rascio & Navari-Izzo 2011). They are outlined – in brief – below.

Defence hypothesis

The toxicity of heavy metals to most organisms is widely-recognised (Nagajyoti et al. 2010) and, arguably, the most intuitively obvious explanation for accumulating high concentrations of toxic molecules (in any taxa) is to use as a defence against being eaten. Thus, the accumulation of extraordinarily high concentrations heavy metals by hyperaccumulators has been proposed as a defence mechanism against herbivory and/or pathogens (Reeves et al. 1981). Despite a myriad of studies on herbivore and pathogen interactions with hyperaccumulating plants, no clear

consensus has been reached (Boyd 2007). Surprisingly, however, it has been clearly and repeatedly demonstrated both that hyperaccumulation can be a deterrent to generalist folivores (Boyd et al. 2002) and that, in other cases, the trait does not reduce or influence levels of herbivory (Huitson & Macnair 2003). It has further been demonstrated that certain specialised arthropod herbivores may preferentially feed on hyperaccumulators; at least in some instances, ironically, to acquire heavy metals for the biosynthesis of toxins to be used in predation-defence (Mesjasz-Przybyłowicz & Przybyłowicz 2001). It has also been demonstrated that hyperaccumulators display increased resistance to plant pathogens such as *Pseudomonas syringae* (Fones et al. 2010). It would appear, therefore, that the role of hyperaccumulation in the selective milieu is a complex one requiring further investigation.

Interference hypothesis

This hypothesis can be further divided into competitive interference[3] and allelopathic interference[4] (Boyd & Martens 1992). The former strategy is intuitively compelling, especially when considered in the context of anthropogenic soil contamination and the many examples of monospecific (hyperaccumulator) colonisation of such sites (Bhargava et al. 2012; Goolsby & Mason 2015; Krämer 2010; Krämer et al. 2007; Sobczyk et al. 2017). Despite the longevity of this idea (Baker & Brooks 1989), little experimental work has been done on intraspecific competition between hyperaccumulators and non-accumulators (Rascio & Navari-Izzo 2011). Similarly, there is little evidence for allelopathy in hyperaccumulators (Morris et al. 2009; Rascio & Navari-Izzo 2011; Zhang et al. 2007). Given the key role of metal ion chelation in minimising phytotoxicity, the shedding of plant material seems unlikely to have a pronounced impact on edaphic concentrations of free, ionic forms of accumulated metals. Furthermore, hyperaccumulation biogeography is canonically associated with serpentine soils; it is hard to envisage the selective utility of "enriching" these soils with chelate-metal complexes.

[3] i.e. competitive advantage on metalliferous soils.
[4] i.e. the shedding of metal-containing tissues to increase metal content of soil in the immediate vicinity of a hyperaccumulator and thus deter growth of competitors.

4. Why study hyperaccumulation?

The intrigue inherent to a trait as extreme as hyperaccumulation is undeniable to anyone with even a passing interest in plant evolution and molecular functioning. There are, however, several very practical reasons for wishing to understand the molecular mechanisms behind hyperaccumulation. The most notable of these relate to potential commercial and/or agronomic utility, specifically phyto-remediation, phyto-mining and the bio-fortification of crop plants.

The contamination of agricultural land with anthropogenic pollutants, including heavy metals, is a widely-reported problem across the globe (Mahar et al. 2016; Tauqeer et al. 2016). An elegant solution to this issue would be to harness the ability of hyperaccumulators to extract edaphic metals. Although the development of soil remediation methodologies is still very much in its infancy (Rascio & Navari-Izzo 2011), there is some experimental support for the practicability of hyperaccumulator-driven soil-remediation with regard to nickel, arsenic, cadmium and palladium (Rascio & Navari-Izzo 2011; Tauqeer et al. 2016).

Given the commercial value of many metals, the corollary to phyto-remediation is phyto-mining; the use of hyperaccumulating plants to extract metal from the soil and the subsequent harvesting of "bio-ore" therefrom (Anderson et al. 1999; Brooks et al. 1999; Rascio & Navari-Izzo 2011). The feasibility of such mining has been demonstrated in *Streptanthus*, *Berkheya*, *Alyssum*, *Brassica* and *Iberis* for a variety of metals including nickel, thallium and even gold (Anderson et al. 1999).

Their purely commercial value notwithstanding, many metals are essential micro-nutrients for our species (Zhu et al. 2016). For example, iron deficiency anaemia has been reported to affect nearly a quarter of the population worldwide (Clemens 2006; Cvitanich et al. 2010). As such, the bio-fortification of crop plants through the molecular manipulation of metal homeostasis networks may have tremendous implications for public health globally – the tools for which may be provided through

a better understanding of the minutiae of hyperaccumulation processes (Clemens et al. 2002; Krämer 2010).

5. The molecular basis of hyperaccumulation

The constitutive overexpression of metal transporter proteins in hyperaccumulator roots has been demonstrated to result in increased zinc uptake in the zinc/cadmium hyperaccumulators *A. halleri* and *N. caeru*lescens (Assunção et al. 2001; Pence et al. 2000). This is considered to be a feature common to all hyperaccumulators (Chaney et al. 1997). The rapid shuttling of metals from the roots to the aboveground tissues is another important feature of hyperaccumulation – and runs contrary to the strategy taken by non-accumulators, where the photosynthetically active shoot tissues are protected from metal toxicity by chelation and storage of metal ions in the roots (Krämer 2010; Rascio & Navari-Izzo 2011; Verbruggen et al. 2009). Although the extent to which hyperaccumulators utilise ligands in the metal translocation process is not yet clear (Haydon & Cobbett 2007; Rascio & Navari-Izzo 2011), the chelation (and subsequent sequestration) of metal ions in leaf vacuoles is another important feature of hyperaccumulation and metal tolerance. Free amino acids, particularly histidine and nicotinamine, play key roles in the complexation of nickel and other bivalent cations (Callahan et al. 2006). Constitutive overexpression of the encoding genes has been demonstrated in several hyperaccumulators (Ingle et al. 2005; Kerkeb & Krämer 2003).

Despite a growing body of literature regarding metal homeostasis and transport in plants, there is a surprising paucity of experimental data on nickel transport proteins and their encoding genes (Merlot et al. 2014; Rascio & Navari-Izzo 2011). In *Arabidopsis*, proteins from the ZIP/IRT[5] and IREG/FPN[6] families have been shown to play a role in the transport of nickel (Morrissey et al. 2009; Nishida et al. 2011; Schaaf et al. 2006; Vert et al. 2002). In *Noccaea*, YSL/OPT[7] and NRAMP[8] family proteins have been similarly implicated (Gendre et al. 2007; Mizuno et al. 2005; Wei et al. 2009).

[5] Zinc-regulated transporter & IRT-like protein / iron-regulated transporter.
[6] Iron-regulated protein / ferroportin.
[7] Yellow stripe-like / oligopeptide transporter.
[8] Natural resistance-associated macrophage protein.

More recently, IREG/FPN and NRAMP proteins have been demonstrated to mediate nickel transport in the nickel hyperaccumulator *Psychotria gabriellae* (Merlot et al. 2014).

Other proteins implicated in the transport of metal ions include the HMA[9], CDF/MTP[10], MFS[11], MATE, VIT and COPT protein families (Krämer et al. 2007; Rascio & Navari-Izzo 2011). It might be noted that protein substrate-specificity and -affinity vary widely across and within protein transporter families. For example, in *A. thaliana*, IRT1 has a low affinity for iron (II) and will readily transport other transition metals, particularly under iron-deficient conditions (Li & Kaplan 1998). Even the high-affinity *Arabidopsis* iron transporter, IREG2, will bind to and transport nickel ions when iron concentrations are minimal (Schaaf et al. 2006). Such results highlight the inadequacy of our current understanding of metal transport processes in plant metal homeostasis.

6. ***Senecio coronatus*: a possible model for nickel hyperaccumulation**

Senecio coronatus (Thunb.) Harv., Asteraceae, is a widespread southern African perennial (Figure 1) which occurs on both serpentine and non-serpentine soils in the region (Mesjasz-Przybyłowicz et al. 1997). The ultramafic rocks of the Barberton Greenstone Belt in Mpumalanga Province, South Africa, have resulted in the genesis of serpentine soils, where edaphic nickel concentrations may exceed 1000 mg per kg of soil dry weight (Morrey et al. 1989). Populations of *S. coronatus* grow on these soils and, although shoot nickel content varies widely, have been demonstrated to accumulate over 4000 mg nickel per kg shoot tissue (dry weight).

Unusually, different *S. coronatus* populations may be comprised of either hyperaccumulating or non-accumulating phenotypes. Even more remarkable is that these phenotypes appear to have a genetic rather than environmental basis. In a recent study, no significant correlation could be drawn between soil nickel content (or pH) and shoot nickel content across different populations and soil types, nor did

[9] Heavy metal transporting ATPases.
[10] Cation diffusion facilitator / metal transporter protein.
[11] Major facilitator superfamily.

soil-swap experiments influence the accumulation phenotype (Meier et al. in review), thus strongly suggesting an entirely genetic basis for nickel accumulation in this species. *S. coronatus* therefore exemplifies a model for investigating the molecular basis of nickel hyperaccumulation, in that direct genomic and transcriptomic comparisons can be made between geographically-proximal hyperaccumulating and non-hyperaccumulating individuals without regard for confounding variables necessarily attendant on even species-level divergence. This combination of features represents a unique opportunity to elucidate aspects of nickel transport, accumulation and tolerance.

Figure 1. *Senecio coronatus* plant from the Barberton region of Mpumalanga Province, South Africa, cultivated under greenhouse conditions at the University of Cape Town.

7. Transcriptomic analysis of *Senecio coronatus*

It has been suggested that modifications to gene expression, rather than the evolution of novel genes, underlie the process of hyperaccumulation (Krämer 2010; Rascio & Navari-Izzo 2011). This hypothesis has been supported in studies where expression patterns have been compared between closely-related hyperaccumulators and non-accumulators (Becher et al. 2004; Van De Mortel et al. 2006; Weber et al. 2004) and the constitutive upregulation of transport-associated genes has been observed in the former. Such studies have relied heavily on the genetic similarities between zinc/cadmium hyperaccumulator *A. halleri* (and, to a lesser extent, *N. caerulescens*) and *A. thaliana* in order to exploit existing molecular tools developed for this species. However, the value of intraspecific comparisons is obvious, as the confounding variables contingent on species-level divergence, genetic drift etc. are minimised (Pollard et al. 2014). Indeed, of the approximately 8000 genes showing differential expression in a comparison between *A. halleri* and *A. thaliana*, only 25 were associated with metal transport (Weber et al. 2004).

In a recent study (Meier et al. in review), the *S. coronatus* transcriptome was assembled *de novo* for plants from both the hyperaccumulating and non-accumulating ecotypes. Nearly 400 genes were identified as being upregulated in the shoot tissue of hyperaccumulating populations, including homologs of *Arabidopsis IRT1* and *IREG2*, both of which have been shown to be involved in nickel transport (Schaaf et al. 2006). A functional enrichment analysis was performed and gene ontology (GO) terms associated with metal transport proteins allowed for the tentative identification of other candidate genes for nickel accumulation in this species. Included in this subset of upregulated genes were genes tentatively identified as encoding proteins in the MATE, VIT and COPT families.

8. The MATE, VIT and COPT protein families

The multidrug and toxic compound extrusion (MATE) protein family is an important and taxonomically widely-distributed group associated with the transport of a broad range of chemical substrates. In plants, these include hormones, secondary

metabolites, organic acids and metal cations (Liu et al. 2016). In *A. thaliana*, as many as 56 putative MATE family transporters have been identified (Li et al. 2002), several of which have been linked to metal detoxification (Diener et al. 2001; Li et al. 2002). In hyperaccumulators, a MATE-group gene (*FRD3*[12]) has been shown to be constitutively overexpressed in *A. halleri* (Talke 2006) and *N. caerulescens* (Van De Mortel et al. 2006). MATE proteins have also been associated with metal detoxification and tolerance in crop plants, including soybean (Liu et al. 2016) and sorghum (Magalhaes et al. 2007). In studies involving the heterologous expression of MATE family proteins in yeast, an *Arabidopsis* MATE protein (*ALF5*[13]) has been shown to increase resistance to a toxic ammonium cation (Diener et al. 2001) and it was suggested that this was due to a direct role in vacuolar sequestration and/or cellular efflux. A homolog of *AtIREG* from the nickel hyperaccumulator *Psychotria gabriella* has similarly been shown to increase nickel tolerance in yeast (Merlot et al. 2014). Furthermore, the expression of PgIREG1 in yeast was demonstrated to reduce the accumulation of intracellular of nickel, suggesting that PgIREG1 acts as a cellular efflux protein rather than a vacuolar cation pump – although this protein was subsequently shown to localise to the tonoplast *in planta*, suggesting that it is responsible for metal transport into the vacuole.

Copper and iron both play essential roles as cofactors in many metabolic pathways (respiration, photosynthesis etc.), yet are toxic if intracellular concentrations are too high (Perea-García et al. 2013). As such, copper/iron homeostasis is tightly regulated in plant cells (Perea-García et al. 2013; Sancenón et al. 2004). The vacuolar iron transporter (VIT) proteins play a critical role in plant cell iron homeostasis, facilitating the shuttling of excess iron (or other divalent cations) into the vacuole for sequestration (Slavic et al. 2016). Initially investigated in yeast (*CCC1*), the *Arabidopsis* ortholog *VIT1* has been shown to be responsible for controlling vascular iron sequestration, which is essential for the minimisation of intracellular iron toxicity (Kim et al. 2006). Furthermore, VIT proteins have previously been shown to rescue the phenotype of iron-sensitive yeast mutants (Gollhofer et al. 2014) – and the iron

[12] Ferric reductase defective 3.
[13] Abberant lateral root formation 5.

concentration in the vacuoles of transformed mutants was 3- to 4-fold higher than that of untransformed yeast cells. A similar pattern was observed with the rice homologs of *AtVIT1* and *AtVIT2*, where heterologous expression in yeast rescued iron-sensitive phenotypes and increased vacuolar iron concentrations (Zhang et al. 2012). Significantly, however, the phenotypes of both zinc- and manganese-sensitive mutants were similarly rescued when these *AtVIT1*/*AtVIT2* homologs were expressed, indicating a broader metal specificity and thus a wider role for VIT proteins in cell metal homeostasis.

Copper homeostasis in plants is largely regulated by a group of high-affinity copper transporter (COPT) proteins although it has been suggested that *COPT* genes may also play a parallel role in iron regulation under certain conditions (Perea-García et al. 2013). In an early study of *Arabidopsis* copper transport mechanisms, *AtCOPT1* expressed in yeast resulted in increased sensitivity to copper (Kampfenkel et al. 1995). Similarly, it has been shown that the heterologous expression of *AtCOPT1* and *AtCOPT2* in yeast mutants deficient in high-affinity copper transporters resulted in phenotypic rescue (Perea-García et al. 2013), thus strongly suggesting it functions as a copper transporter. Furthermore, it was shown that *AtCOPT1* and *AtCOPT2* were upregulated in response to both copper and iron deficiencies in *Arabidopsis*, indicating that COPT proteins may be involved in the transport of other metals. It has since been confirmed that the regulation of COPT proteins may be driven by iron, manganese and zinc concentrations (Yuan et al. 2011).

MATE proteins have been shown to localise to the plant plasma membrane (Magalhaes et al. 2007) and tonoplast (Shitan et al. 2014) and are widely-regarded as having 12 transmembrane domains (Braibant et al. 2002; Brown et al. 1999; Maron et al. 2010; Shitan et al. 2014).VIT proteins have been demonstrated to localise in the tonoplast in members of Brassicaceae (Kim et al. 2006; Zhu et al. 2016) and Poaceae (Zhang et al. 2012) and are predicted to have 5 transmembrane domains (Slavic et al. 2016; Zhang et al. 2012). COPT proteins have been shown to localise to the plasma membrane in *Arabidopsis* (Perea-García et al. 2013) and *Oryza* (Yuan et al. 2011). It

has been suggested that COPT proteins contain 3 transmembrane domains (Puig & Thiele 2002).

9. Project aims and objectives

Broadly speaking, the aim of this project was to determine whether three proteins shown to be overexpressed in hyperaccumulating populations of *S. coronatus* play a role in nickel transport. Elucidating the function of these proteins will add to our growing, but as yet inadequate, understanding of the molecular mechanisms behind nickel transport and accumulation.

The objectives for the project were threefold. Initially, we sought to validate the nucleotide sequence data for the three genes generated through prior RNA-Seq work (by which the upregulated genes were identified); the validation of which allowed for phylogenetic and bioinformatic analyses to proceed. This was particularly with respect to similarities in the amino acid sequences encoded by these genes and the amino acid sequences of related proteins in other taxa. Such a strategy further allowed comparison of amino acid sequences and protein topologies with known metal transporters in other, better-studied species.

Secondly, we wished to investigate the role of the proteins encoded by these genes in nickel transport. More specifically, the heterologous expression of these proteins in a metal-sensitive yeast strain was used to assess the capacity of these proteins to transport yeast, either into, or out of, yeast cells.

The third and final objective was to establish the subcellular localisation of these proteins. In order to do so, YFP-tagged fusion proteins were generated and expressed in living plant cells and fluorescence microscopy utilised to record fluorescent signal. Establishing whether these proteins are capable of transporting nickel, combined with an understanding of where in the plant cell they target, represents an important first step in understanding the mechanisms behind nickel transport and accumulation in this species.

II METHODS & MATERIALS

10. Plant material

Plant collection

Plant material was originally collected in 2012 from the Barberton Greenstone Belt, Mpumalanga Province, Republic of South Africa (Figure 2). Whole plants were collected from a total of 15 sites, from both serpentine and non-serpentine soils. The status of each plant as a hyperaccumulator or non-accumulator was determined in the field[14] through the use of dimethylglyoxime-impregnated (1% w/v) filter paper (Reeves et al. 1996). Both soil types were collected for the purposes of *ex situ* plant cultivation.

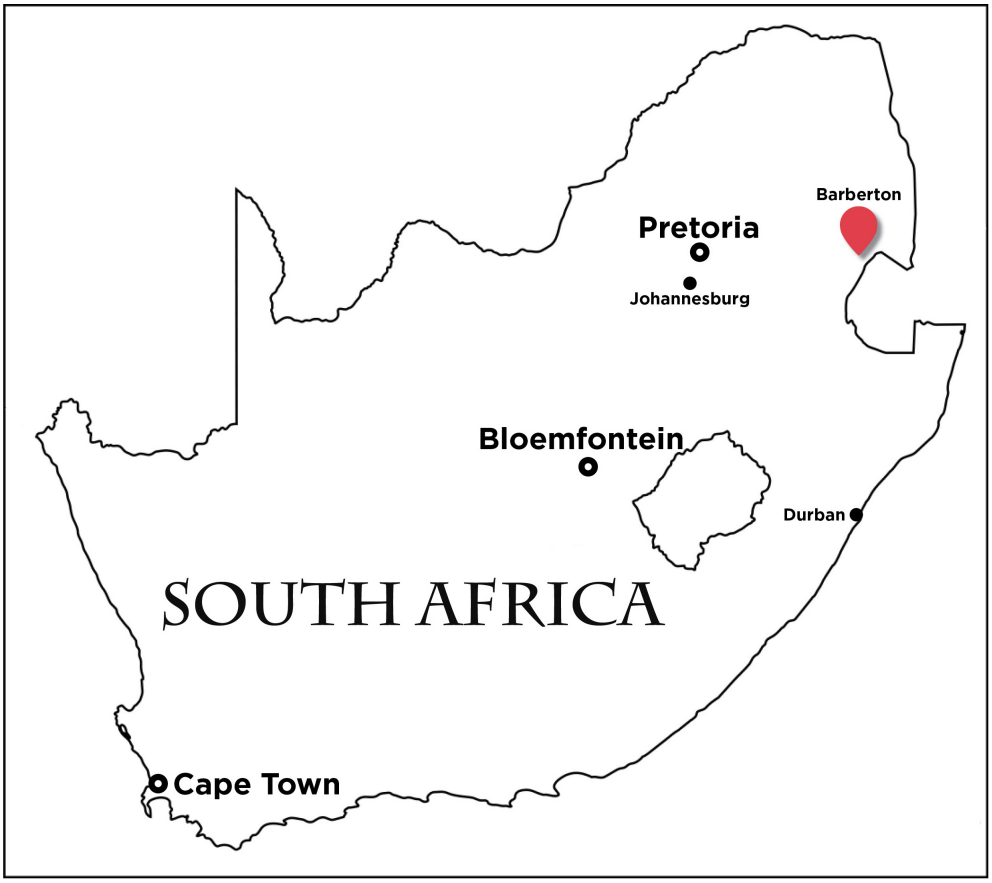


Figure 2. Map of the Republic of South Africa indicating (red) the town of Barberton in Mpumalanga Province. Plants were collected from the Barberton Greenstone Belt in the vicinity of Barberton.

[14] Field determinations were later confirmed through the use of Inductively Coupled Plasma Mass Spectrometry (Meier et al. in review).

Plant cultivation

Initially, each plant was transferred to a pot containing soil from its individual collection site. The plants used for RNA-Seq were potted in a 1:1 mix of serpentine and non-serpentine soil. The plants were cultivated in an unheated greenhouse under natural light conditions. The watering regime was changed according to season and day-to-day temperature fluctuations. Opportunistic growth of other plant species and bryophytes was managed manually. Phytophagous insects were controlled using a 0.5% (v/v) solution of the mercaptothion-based insecticide, Malasol[15].

Tissue harvesting

Tissue harvesting was performed in the most non-invasive way possible, so as to limit damage to plants and subsequent morbidity. Generally, no more than one medium-sized leaf was harvested from each plant.

11. Microbial strains

Escherichia coli

DH5α is a non-pathogenic *E. coli* strain optimised for transformation efficiency. It is sensitive to the ccdB protein and thus suitable for cloning experiments involving vectors containing the *ccdB* gene within the multiple cloning site ("MCS"). One Shot® Chemically Competent *E. coli*[16] ("one shot") are a commercially-available *E. coli* strain with a particularly high transformation efficiency and are recommended for the propagation of Gateway Vectors. Both these strains were cultured in either liquid LB or on LB agar[17] plates (1% w/v), supplemented with the relevant antibiotic in the case of plasmid-transformed cells. Cultures were typically incubated (with shaking of approximately 250 RPM in the case of liquid cultures) for 12-16 hours at 37°C.

[15] Efekto, Bryanston, South Africa.
[16] Life Technologies, Carlsbad, USA.
[17] Bacteriological grade agar used throughout.

Agrobacterium tumefaciens

GV3101 (Holsters et al. 1980) is a commonly-used non-oncogenic *A. tumefaciens* strain. It was cultured in either liquid LB or on LB agar plates (1% w/v) containing rifampicin (100 µg/mL) and gentamycin (15 µg/mL), additionally supplemented with plasmid-specific antibiotics where relevant.

Saccharomyces cerevisiae

The *S. cerevisiae ZHY3* strain (Zhao & Eide 1996) is a *zrt1zrt2*[18] mutant and thus deficient in functional zinc-uptake transporters. Furthermore, this strain is an uracil auxotroph (*ura3* mutant) allowing for selection of positive transformants. This yeast strain was donated by Sylvain Merlot[19]. It was cultured in either liquid YNB or on YNB agar plates (2% w/v).

12. Plasmids

pENTR1A

Gateway® pENTR™ 1A Dual Selection Vector[20] ("pENTR1A") is a commercially-available Gateway-compatible entry vector. It contains both the *ccdB* "suicide gene" (which allows for negative selection in *E. coli* strains sensitive to the ccdB toxin) and a kanamycin-resistance gene (which allows positive selection of transformants). This plasmid was propagated from laboratory-maintained *E. coli* glycerol stocks.

pDR195gtw

The pDR195gtw plasmid ("pDR195") is a variant of the pDR195 yeast expression vector modified to be Gateway-compatible (Oomen et al. 2009). It contains an ampicillin-resistance gene to facilitate propagation in *E. coli*. The pPMA1 promoter drives constitutive transcription in *S. cerevisiae*. It was donated by Merlot Sylvain[21].

[18] Zinc-regulated transporter.
[19] Institut des Sciences du Végétal, France.
[20] Life Technologies, Carlsbad, USA.
[21] Institut des Sciences du Végétal, France.

pEarleyGate101

The pEarleyGate101 plasmid ("pG101") is a Gateway-compatible plant expression vector. It contains a kanamycin-resistance gene for selection in bacteria and a BASTA-resistance gene for selection in plants. It is engineered to allow in-frame production of C-terminal-*EYFP* fusion proteins in plant cells. Transcription is driven by the promoter *CaMV 35S* promoter. It was donated by Laura Roden[22].

13. RNA manipulation

RNA extraction

A modified version of the acid guanidinium thiocyanate-phenol-chloroform extraction method (Chomczynski & Sacchi 1987) was used to extract total RNA from *S. coronatus* leaf tissue. See Table 1 for a list of reagent constituents. Tissue samples of approximately 100 mm^2 were collected from live plants and immediately flash-frozen in liquid nitrogen. Each tissue sample was placed into a 1.5 mL microfuge tube containing three 2 mm stainless steel ball bearings, after which 1 mL of RNA extraction solution was added to each tube. The tissue samples were then homogenised in a commercial paint shaker for approximately 15 minutes. The samples were centrifuged in a table-top centrifuge at 12000 RPM for 5 minutes at 4°C. Being careful not to disturb the pellet of cell debris, approximately 600 µL of the supernatant was removed and added to a microfuge tube containing 200 µL chloroform. The tubes were inverted, vortexed for 30 seconds, then incubated for 5 minutes at room temperature. After briefly vortexing again, the samples were centrifuged at 12000 RPM for 15 minutes at 4°C. Approximately 200 µL of the aqueous phase was carefully aspirated and added to a microfuge tube containing an equal volume of RNA precipitation solution (see Table 1). The tubes were incubated for 10 minutes at room temperature and then centrifuged at 12000 RPM for 15 minutes at 4°C. After visually establishing the presence and location of the RNA pellet, the supernatant was discarded. The pellet was then washed in ethanol (75% v/v, diluted with DEPC-treated water), briefly vortexed at low speed, and centrifuged at 6000 RPM for 5

[22] University of Cape Town, South Africa.

minutes at 4°C. The ethanol was removed and the RNA pellet allowed to air-dry at room temperature. The pellet was then suspended in 22 µL of DEPC-treated Millipore water and incubated on a heating block for 10 minutes 55°C. When fully re-suspended, the sample was centrifuged at 5000 RPM for 5 minutes at room temperature, after which the RNA-containing supernatant was transferred to a clean tube and stored at -80°C.

Analysis of RNA quality

RNA concentration and purity was evaluated using a NanoDrop® ND-1000 Spectrophotometer[23] with particular consideration being given to the 260 nm / 280 nm ratio as an indication of guanidium/carbohydrate contamination. RNA quality was further assessed through electrophoresis, using a formaldehyde-based RNA denaturing gel. The gel contained 1 x MOPS buffer, agarose (1% w/v) and formaldehyde (37% v/v). The volume of 2 µg of RNA was calculated (as according to the NanoDrop readings), transferred to a clean microfuge tube, and adjusted to a total volume of 10 µL with DEPC-treated Millipore water. To this was added a one-fifth volume of RNA loading buffer, then the samples were denatured by incubating on a heating block for 5 minutes at 65°C and then plunging into an ice bath before loading onto the gel. A 1x MOPS buffer solution was used as a running buffer. The RNA gel was visualised on a Syngene G:Box gel doc system[24].

14. cDNA preparation

cDNA synthesis

cDNA was synthesised using SuperScript™ III Reverse Transcriptase[25] from 1 µg RNA. The protocol supplied by the manufacturer was followed, except for a reduction from 1 µL to 0.5 µL of transcriptase used per reaction.

[23] NanoDrop Products, Wilmington, USA.
[24] Synoptics Group, Cambridge, UK.
[25] Life Technologies, Carlsbad, USA.

Analysis of cDNA quality

The integrity of the resulting cDNA was tested by performing a PCR using *IRT1* primers (Table 2) – a procedure that had been repeated successfully on *S. coronatus* cDNA on numerous prior occasions. The PCR utilised Super-Therm *Taq* Polymerase[26]. Two reactions each of 20 µL were performed, following the manufacturer's recommendations with regard to reagent volumes and concentrations. An annealing temperature of 55°C was used and 35 x melt/anneal/elongate cycles were performed using a Gene Amp PCR System 2700[27]. The PCR product was electrophoresed in an agarose gel (1% v/v) made with TAE buffer. Ethidium bromide (0.5 µg/mL) was added to the molten gel before it was allowed to set. The electrophoresis was run for approximately an hour at 70-80V, using 1x TAE as a running buffer. The gel was visualised on a Syngene G:Box gel doc system [28].

15. DNA manipulation

Preparation of insert DNA

PCR amplification of all GOIs was initially carried out using Super-Therm *Taq* Polymerase[29] or the KAPA *Taq* PCR Kit[30]. After confirming the presence of amplicons of the predicted size through electrophoresis (and, in some cases, sequence analysis), amplification of the GOIs was repeated using high-fidelity polymerases, either the KAPA HiFi HotStart ReadyMix PCR Kit[31] or Q5® High-Fidelity DNA Polymerase[32]. Manufacturers' recommendations were followed with respect to reagent volumes/concentrations. An initial denaturation phase (5 minutes at 94°C) was followed by 35 melt/anneal/elongate cycles and a final elongation phase (10 minutes at 72°C). The cycle melt phase was 15 seconds at 94°C. The annealing phase was 30 seconds – see Table 2 for annealing temperatures. The cycle elongation

[26] Separations Scientific, Honeydew, South Africa.
[27] Applied Biosystems, Foster City, USA.
[28] Synoptics Group, Cambridge, UK.
[29] Separations Scientific, Honeydew, South Africa.
[30] Kapa Biosystems, Wilmington, USA.
[31] Kapa Biosystems, Wilmington, USA.
[32] New England Biosystems, Ipswich, USA

phase was typically 15-30 seconds at 72°C. All PCRs were performed on a GeneAmp® PCR System 2700[33].

PCR products were electrophoresed in agarose gels (1% v/v) and viewed on a UV transilluminator (365 nm). After confirming that band sizes matched those of the expected amplicons, bands were carefully excised with a scalpel blade and transferred to microfuge tubes. The Wizard® SV Gel and PCR Clean-Up System[34] was used to purify DNA, following the manufacturer's "DNA Purification by Centrifugation Quick Protocol". However, DNA was eluted in 30 µL water rather than the recommended 50 µL.

The purified DNA was digested using one or more New England Biolabs restriction enzymes, as per the manufacturer's instructions, with one modification: digestions were incubated for 90 minutes at 37°C. Finally, the digested insert was purified, again using the Wizard® SV Gel and PCR Clean-Up System[35].

Preparation of entry vector

Cells from an existing *E. coli* (one-shot) glycerol stock (containing pENTR1A) were used to inoculate 5 mL LB and incubated for 12-18 hours at 37°C. The PureYield™ Plasmid Miniprep System[36] was used to extract purified plasmids from the liquid culture. The manufacturer's "DNA Purification by Centrifugation Quick Protocol" was followed, using the alternative protocol for larger culture volumes. The concentration of the plasmid purified was quantified using the NanoDrop ND-1000.

The purified plasmid was digested using one or more New England Biolabs restriction enzymes, as per the manufacturer's instructions (but incubated for 90 minutes at 37°C). The digested "backbone" was then electrophoresed in an agarose gel (1% v/v) and viewed on a UV transilluminator (365 nm). After confirming that band sizes matched those expected for the backbone and any fragments, the backbone

[33] Applied Biosystems, Foster City, USA.
[34] Promega, Madison, USA.
[35] Promega, Madison, USA.
[36] Promega, Madison, USA.

was carefully excised with a scalpel blade and transferred to a microfuge tube. The Wizard® SV Gel and PCR Clean-Up System[37] was used to purify the DNA, following the manufacturer's "DNA Purification by Centrifugation Quick Protocol". However, DNA was eluted in 30 µL rather than the recommended 50 µL.

All vector backbones were dephosphorylated prior to ligation- or Gateway-cloning. Shrimp Alkaline Phosphatase[38] ("rSAP") was used and the manufacturer's instructions were followed exactly.

Preparation expression vectors

Expression vectors were initially propagated in liquid *E. coli* (one-shot) cultures. The PureYield™ Plasmid Miniprep System[39] was used to extract purified plasmids. The manufacturer's "DNA Purification by Centrifugation Quick Protocol" was followed, using the alternative protocol for larger culture volumes. The concentration of the plasmid purified was quantified using the NanoDrop ND-1000.

Cloning

Plasmid-insert ligations were carried out using T4 DNA Ligase[40], following the manufacturer-supplied protocol and using a ratio of 1:3 (vector to insert). Reactions were incubated for 1 hour at room temperature.

Gateway cloning was performed using the Gateway™ LR Clonase™ II Enzyme Mix[41]. The manufacturer's instructions were followed exactly, using a ratio of 1:1 (entry clone to destination vector). In the case of pG101 containing a GOI, the plasmid was linearized using a restriction enzyme prior to cloning.

PCR products were electrophoresed on a Power Pac 200[42] in TAE-agarose gels (1% v/v) submerged in TAE buffer. Electrophoresis was performed for approximately

[37] Promega, Madison, USA.
[38] New England Biolabs, Ipswich, USA.
[39] Promega, Madison, USA.
[40] New England Biolabs, Ipswich, USA.
[41] Life Technologies, Carlsbad, USA.
[42] Bio-Rad Laboratories, Hercules, USA.

1 hour at 75-100V, depending on the size and thickness of the gel concerned. The volume of PCR product loaded depended on both its purpose and the size of the gel, but was typically between 22 and 40 µL. PCR product was run alongside one or more lanes loaded with 10-16 µL DNA ladder, generally either GeneRuler DNA Ladder[43] or Quick-Load DNA Ladder[44] in both/either 1 kb and 100 bp ladders.

16. Microbial manipulation

Escherichia coli competent cell preparation

Cells from an existing *E. coli* (DH5α) glycerol stock stored at -80°C were streaked onto a single LB plate (1% w/v agar). This was incubated for 12-18 hours at 37°C. A single colony from the plate was used to inoculate 5 mL LB (no antibiotic), which was incubated with shaking for 12-18 hours at 37°C. The entire inoculum was then added to 100 mL LB, which was incubated with shaking for 2-3 hours at 37°C until the OD_{600} reached approximately 0.4. The culture was immersed in an ice bath for 10 minutes and then aliquoted into pre-cooled centrifuge tubes and spun in an Avanti® J-E Centrifuge[45] using a JA-20 rotor at 4700 RPM for 10 minutes at 4°C. The supernatant was discarded and the cell pellet re-suspended in 1.6 mL ice-cold $MgCl_2$-$CaCl_2$ solution (80 mM $MgCl_2$, 20 mM $CaCl_2$), before being incubated in an ice bath for 30 minutes. The cells were then spun at 4700 RPM for 10 minutes at 4°C and, after discarding the supernatant, re-suspended in 1.6 mL ice-cold $CaCl_2$ solution (100 mM). The cells were incubated in an ice bath for a further 20 minutes, after which 0.5 mL ice-cold glycerol (80% v/v) was added. After thoroughly mixing the solution, aliquots of 100 µL were transferred into pre-cooled microfuge tubes and then flash frozen in liquid nitrogen, before being transferred to -80°C for storage.

Escherichia coli competent cell transformation

Competent cells were removed from storage at -80°C and allowed to defrost on ice. Each ligation/Gateway reaction was spun briefly on a table-top centrifuge, after which 3 µL was added to a pre-cooled microfuge tube on ice. After gently mixing,

[43] Life Technologies, Carlsbad, USA.
[44] New England Biolabs, Ipswich, USA.
[45] Beckman Coulter, Brea, USA.

50 µL of competent cells were added to the tube. The tube was gently flicked and then incubated on ice for 20 minutes. The cells were then heat-shocked in a water bath for 45-50 seconds at 42°C, before being returned to ice for a further 2 minutes. 950 µL LB was added to the tube which was incubated with gentle shaking for 1.5 hours at 37°C. The cells were then spun on a table-top centrifuge at 5000 RPM for 5 minutes. Approximately 900 µL of the supernatant was discarded, the cell pellet re-suspended in the remaining culture, and then spread onto LB plate (1% w/v agar, containing the appropriate selective antibiotic). The plate was then incubated for 12-18 hours at 37°C.

Agrobacterium tumefaciens competent cell preparation

Cells from an existing *A. tumefaciens* (GV3101) glycerol stock stored at -80°C were streaked onto a single LB plate (1% w/v agar, supplemented with 100 µg/mL rifampicin and 15 µg/mL gentamycin). This was incubated for 24-36 hours at 28°C. A single colony from the plate was used to inoculate 10 mL LB (supplemented with 100 µg/mL rifampicin and 15 µg/mL gentamycin), which was then incubated with shaking (approximately 250 RPM) for 24-36 hours at 28°C until the OD_{600} was between 0.5 and 1.0. The culture was then aliquoted into centrifuge tubes and spun in an Avanti® J-E Centrifuge[46] at 3000 RPM for 10 minutes at 4°C. The supernatant was discarded and the cells re-suspended in 0.5 mL ice-cold $CaCl_2$ solution (20 mM) and then aliquots of 100 µL were transferred into pre-cooled microfuge tubes. Cells not used immediately were flash-frozen in liquid nitrogen, before being moved to -80°C for storage.

Agrobacterium tumefaciens competent cell transformation

In the case of freshly-prepared competent cells, 1 µg plasmid DNA was added to a 100 µL aliquot of cells, flicked to mix, and then flash-frozen in liquid nitrogen. In the case of competent cell stored at -80°C, the cells were defrosted on ice for 10-15 minutes, before adding plasmid DNA and flash-freezing in liquid nitrogen. The frozen cells, now mixed with plasmid DNA, were then thawed at 37°C for 5-10 minutes. Then

[46] Beckman Coulter, Brea, USA.

0.9 mL LB was added to each tube, before incubating with gentle shaking for 3 hours at 28°C. The cells were then spun on a table-top centrifuge at 2000 RPM for 3 minutes. After discarding 0.9 mL supernatant, the cell pellet was re-suspended and spread on LB agar plates (1% w/v, supplemented with 100 µg/mL rifampicin, 15 µg/mL gentamycin and 50 µg/mL kanamycin). Finally, the plates were incubated at 28°C for 36-48 hours.

Saccharomyces cerevisiae competent cell preparation

See Table 1 for particulars of solutions used in this section. Cells from an existing *S. cerevisiae* (*ZHY3* strain) glycerol stock stored at -80°C were streaked onto a single YNB plate (2% w/v agar). This was incubated for 24-36 hours at 30°C. A single colony from the plate was used to inoculate 100 mL YPD liquid media, which was incubated with shaking for approximately 24 hours at 30°C until the OD_{600} was between 0.8 and 1.6. The culture was then aliquoted into microfuge tubes and spun in a table-top centrifuge at 3000 RPM for 10 minutes at room temperature. After discarding the supernatant, the cells were suspended in 10 mL TEL buffer and shaken at 30°C for approximately 12 hours. The cells were then pelleted (spun on a table-top centrifuge at 3000 RPM for 10 minutes at room temperature) and re-suspended in 1 mL TEL buffer before being aliquoted into microfuge tubes (100 µL per tube). Cells not used immediately were stored at 4°C for no more than 3 weeks, after which they were discarded.

Saccharomyces cerevisiae competent cell transformation

Salmon sperm DNA was gently heated in order to reduce its viscosity, before 5 µL (approximately 50 µg) was added to aliquots of competent cells. The mixture was flicked in order to mix and then 5 µL plasmid DNA was added. The cells were incubated for 30 minutes at room temperature, after which 0.7 mL PEG-TEL was added and the cells re-suspended therein. The cells were incubated for a further 60 minutes at room temperature and then heat shocked in a 42°C water bath for 5-10 minutes. The cells were spun briefly in a table-top centrifuge, after which the supernatant was discarded and the cells re-suspended in 100 µL TE buffer and then plated on YNB/CSM-uracil agar plates (2% v/v).

17. Diagnostic procedures

Colony PCR

Colony PCRs were used to determine the presence or absence of a GOI in all putative positive transformants. Bacterial colony PCRs were performed according to the parameters for the relevant amplicon (using either gene- or plasmid-specific primers), although with an extended initial denaturation time of 5 minutes at 98°C. A small quantity of bacteria was collected on a 200 µL pipette tip and added to each PCR tube, which was vortexed to thoroughly mix the cells with the PCR reagents before cycling began. Yeast colony PCRs required an additional step; the digestion of yeast cell walls prior to PCR. A whole single yeast colony was collected on a 200 µL pipette tip and added to 20 µL of sodium hydroxide solution (40 mM). The cells were carefully re-suspended and then incubated for 45 minutes at 95°C. The digested yeast cells were centrifuged briefly before 1 µL was added to the PCR reagents prior to initiating the PCR.

Diagnostic digest

The bacterial colonies sampled by colony PCR were used to inoculate liquid cultures which were incubated overnight and used for plasmid extraction. Purified plasmid was digested with one or more restriction enzymes and the amplicon sizes analysed to determine the presence or absence of plasmid and insert.

Plasmid PCR

Where the results from colony PCR and/or diagnostic digestion were ambiguous, plasmid PCR was performed to confirm the presence or absence of the GOI. The extracted plasmid was diluted (10% v/v) and used as template in a standard PCR using gene- or vector-specific primers.

DNA sequencing and analysis

Purified plasmid samples were sent to the Central Analytical Facilities ("CAF") at Stellenbosch University for sequencing on an ABI3730xl DNA Analyser[47]. Sequence data was analysed using 4Peaks Version 1.8[48], MEGA7 (Kumar *et al*. 2015) and BLAST[49].

Glycerol stock preparation

Glycerol stocks were prepared for all microbial strains and transformants. Undiluted glycerol was autoclaved, then diluted with autoclaved Millipore water to 50% v/v. Liquid culture (500 µL) containing live cells was added to a clean microfuge tube. To this was added the glycerol (50% v/v), resulting in a final glycerol dilution of 25% v/v. The tubes were then flash-frozen in liquid nitrogen and stored at -80°C. For all yeast transformants, duplicate glycerol stocks were prepared with both 50% and 30% (v/v) glycerol.

18. Nickel sensitivity assays

An initial assay was conducted using *S. cerevisiae* transformed with empty pDR195 in order to establish a spectrum of nickel toxicity. Three confirmed yeast transformants were cultured overnight and then washed in autoclaved Millipore water before being diluted in water to an OD_{600} of 1. Each of the three cultures was then subjected to three iterations of serial dilution (i.e. OD_{600} of 1, 10^{-1}, 10^{-2} and 10^{-3}). 5 µL of each of the three cultures (for all four dilutions) was spotted onto a series of YNB/CSM-uracil agar plates (2% w/v, buffered to pH 5.5 with MES-KOH) containing increasing concentrations of nickel. The plates were incubated for approximately 48 hours at 30°C before yeast colony growth was scored across all concentrations. Having established the concentration at which nickel began to have noticeable deleterious effects on yeast colony growth, the assay was repeated with pDR195-*ScMATE*, pDR195-*ScVIT*, pDR195-*ScCOP*, pDR195-*AtIREG2* as well as the empty pDR195 plasmid.

[47] Applied Biosystems, Foster City, USA.
[48] Nucleobytes, Amsterdam, Holland.
[49] https://blast.ncbi.nlm.nih.gov/blast.cgi

19. Protein subcellular localization

Infiltration liquid

A recently-devised novel method for transient gene-expression in plants (Xu et al. 2014) was followed, with slight modifications. For each *Agrobacterium* transformant to be used in the transfection, 100 mL of infiltration liquid was made. The infiltration liquid comprised of the following reagents: D-glucose (41.65 mM), $CaCl_2$ (100 mM), MES-KOH (100 mM, pH 5.6), 6-benzylaminopurine (0.011 uM), Silwet L-77[50] (0.01% v/v), $MgCl_2$ and acetosyringone (12.5 mM, dissolved in dimethylformamide or dimethyl sulfoxide).

Onion pre-treatment

A locally-cultivated variety of red onion was used in this experiment. Only medium to large bulbs of similar appearance were selected. Pre-treatment consisted of incubation in darkness for 48-72 hours at 28°C.

Infiltration procedure

A. tumefaciens glycerol stocks containing the pG101/GOI plasmids were used to inoculate 50 mL liquid LB (supplemented with 100 µg/mL rifampicin, 15 µg/mL gentamycin and 50 µg/mL kanamycin). These cultures were incubated with vigorous shaking for 24-36 hours at 28°C. When the OD_{600} reached approximately 1.5, 25 mL of the inoculum was collected and centrifuged in an Avanti® J-E Centrifuge[51] using a JA-14 rotor at 5000 RPM for 10 minutes at 18°C. The pelleted cells were gently re-suspended in 20 mL infiltration liquid (at room temperature). The centrifugation/re-suspension process was repeated 4 times. Following the final re-suspension, the cells were diluted with infiltration liquid to $OD_{600} = 0.1$. A scalpel blade was used to carefully remove the bulb tunic of the onion. Two parallel longitudinal cuts were then made, allowing the 3 or 4 outermost scales to be folded down. A narrow gauge insulin needle and syringe was then used to inject 50-150 µL infiltration liquid between the adaxial epidermis and the mesophyll (Figure 3). The scales and tunic were then

[50] OSi Specialities, Danbury, USA.
[51] Beckman Coulter, Brea, USA.

returned to their original positions and held in place with adhesive tape, before the onions were returned to the dark and incubated for 72 hours at 28°C.

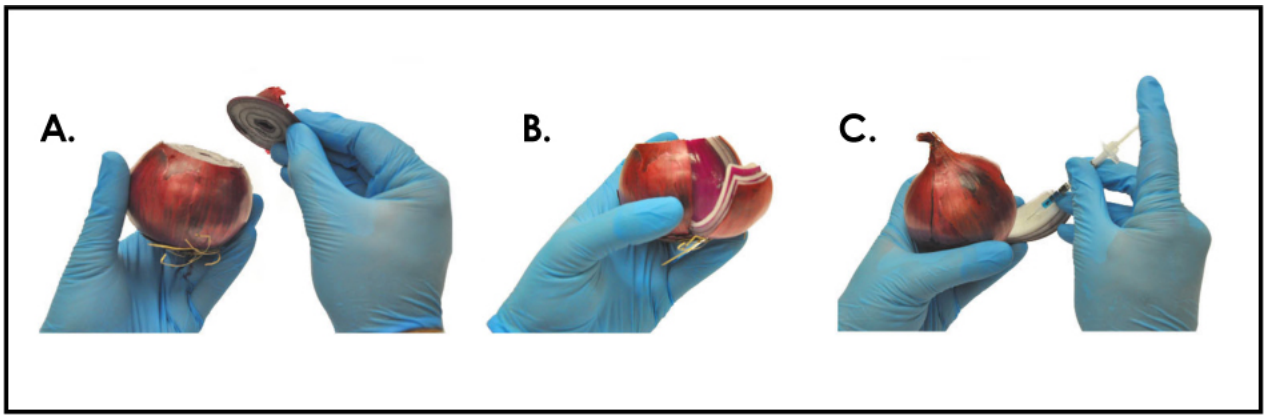


Figure 3. Physical steps in the infiltration of living onion epidermal cells during *Agrobacterium*-mediated *in planta* transient transfection. A. Removal of the bulb tunic. B. Peeling open the outer scales. C. Injection of infiltration liquid into the interface of the adaxial epidermis and mesophyll.

Fluorescence microscopy

Sections of adaxial epidermis (approximately 20 mm^2) in close proximity to the infiltration site were carefully peeled from the onion scales. Each section was placed onto a 10-20 µL droplet of water pipetted onto the surface of a standard microscope slide. A coverslip was carefully lowered onto the epidermis/droplet, creating a wet mount slide.

Onion cells were visualised using the Nikon Eclipse Ti-E Inverted Microscope System[52], linked to the proprietary NIS-Elements AR imaging software. The cells were subjected to laser excitation of 490-510 nm and emitted light was collected between 520-550 nm at both 4x and 10x magnifications. Bright field images were taken both 4x and 10x magnifications.

[52] Nikon Corporation, Tokyo, Japan.

Table 1. Solutions used in study, including concentration (C) and pH.

Solution	Reagents	C	pH
RNA manipulation			
RNA extraction solution	Sodium acetate (0.1 M, pH 5.2) Guanidine thiocyanate (0.8 M) Ammonium thiocyanate (0.4 M) Glycerol (5% v/v) Phenol (38% v/v, pH 4)	-	-
RNA precipitation solution	Isopropanol (1x volume) Sodium chloride (1.2 M, 0.5x volume) Sodium citrate (0.8 M, 0.5x volume) [Made with DEPC treated water]	-	-
MOPS buffer	3-(*N*-morpholino)propanesulphonic acid (200 mM) Ethylenediaminetetraacetic acid (10 mM) Sodium acetate (50 mM, pH 7 with NaOH) [Made with DEPC treated water]	1x	7
RNA loading buffer	Glycerol (50% v/v) EDTA (1 mM) Bromophenol blue (0.4% v/v) Ethidium bromide (50µg/mL)	-	-
Yeast competent cell preparation			
EDTA	Ethylenediaminetetraacetic acid	1 M	8
Tris	Tris(hydroxymethyl)aminomethane	1 M	8
PEG	Polyethylene glycol	50% w/v	-
TEL	Lithium acetate (0.5 M) Tris (50 mM) EDTA (0.5 M)	5X	-
TEL buffer	Lithium acetate (1 M) Tris (1 M, pH 8) EDTA (0.5 M)	0.1 M	-
PEG-TEL	PEG 4000 TEL	40% v/v	-
Miscellaneous			
LB	Tryptone powder (1% w/v) Yeast extract (0.5% w/v) Sodium chloride (0.5% w/v)	-	7
TAE buffer	Tris(hydroxymethyl)aminomethane (40 mM) Ethylenediaminetetraacetic acid (1 mM) Glacial acetic acid (0.11% v/v)	1x	-
YNB	Yeast nitrogen base (Sigma Y0626) (0.67% w/v) Dextrose (0.5 w/v)	-	-
YNB/CSM-uracil	Yeast nitrogen base (Sigma Y0626) (0.67% w/v) CSM-uracil (Sigma Y1501) (0.2% w/v) Dextrose (0.5% w/v)	-	-
YPD	Yeast extract (1% w/v) Peptone powder (2% w/v) Dextrose (2% w/v)	-	-

Table 2. Primer sequences and number of nucleotides, annealing temperatures, and expected amplicon sizes of PCR primers used in experimental work. Restriction enzyme sites indicated in bold.

Primer	Sequence	n	T_A (°C)	E.A.S.*
Primary Cloning				
MATE-F	CTGA**GGATCC**ATGGCCATCAATGGCACTGG	30	56	1491
MATE-R	CTGA**GAATTC**TCAATGAACAACTGCCCATTTCT	33		
VIT-F	CTGA**GGATCC**ATGGCCGGAAATATTGGTGG	30	60	801
VIT-R	CTGA**GAATTC**CTAATCTTCCAAACCAGCCACA	32		
COP-F	CTGA**GTCGAC**ATGATGCATATGACTTTTTGTTGG	34	63	507
COP-R	CTGA**GAATTC**TTATAACAAGCAAACTAACACACT	34		
Secondary Cloning				
MATE-F (2)	CTGA**GGATCC**ATGGCCATCAATGGCACTGG	30	57.6	±1500
MGFP-R	CTGA**GAATTC**TTATGAACAACTGCCCATTTCT	32		
VIT-F (2)	CTGA**GGATCC**ATGGCCGGAAATATTGGTGG	30	59.5	±800
VFFP-R	CTGA**GAATTC**TTATCTTCCAAACCAGCCACA	31		
Miscellaneous				
IRT1-F	AC**GGATCC**ATGGCTTCAAGTTCAAAAAATGTCA	33	61	±1050
IRT1-R	AC**CTCGAG**TTAAGCCCATTTTGCCATCAG	29		
pDR AMP F	ATGAGTATTCAACATTTCCGTGTC	24	53	861
pDR AMP R	TTACCAATGCTTAATCAGTGAGGC	24		

*Expected amplicon size

III RESULTS & DISCUSSION

1. PCR amplification of full-length *MATE*, *VIT* and *COP* cDNAs

Prior to the initiation of any experimental work involving the expression of the MATE, VIT and COP proteins in live cells, two preliminary objectives needed to be met. First, the RNA-Seq derived nucleotide sequences for *MATE*, *VIT* and *COP* required validation. This is discussed in the next section. Second, and more practically, the in-frame cloning of *MATE*, *VIT* and *COP* into an expression vector was needed in order to facilitate the experimental expression work done in the two latter sections of this chapter. This initial cloning work and related preliminaries are described below.

RNA extraction and synthesis of *S. coronatus* cDNA

Extraction of total RNA from *S. coronatus* tissue proved more challenging than the equivalent procedure in the model *A. thaliana*, as the tougher, thicker *S. coronatus* leaf tissue was more difficult to fully homogenise. As such, the RNA yields were less than would be expected for *A. thaliana*. Furthermore, the 260 nm / 280 nm ratio of extracted RNA was generally low (seldom exceeding 0.6), usually indicative of protein contamination. Attempts to optimise the extraction method and/or to column-purify the extracted RNA did not result in any observable improvements to this ratio. Similarly, the concentration of extracted RNA was lower than that typically attained in RNA extractions using *A. thaliana* tissue, seldom exceeding 700 ng/µL. These problems notwithstanding, electrophoresis of extracted RNA (Figure 4) revealed sharply-defined ribosomal RNA bands, typically indicative of acceptable sample integrity. cDNA synthesised from the RNA samples was tested using *IRT1* primers and clear bands of the expected size were observed after electrophoresis (Figure 5), indicating that the cDNA synthesis was successful.

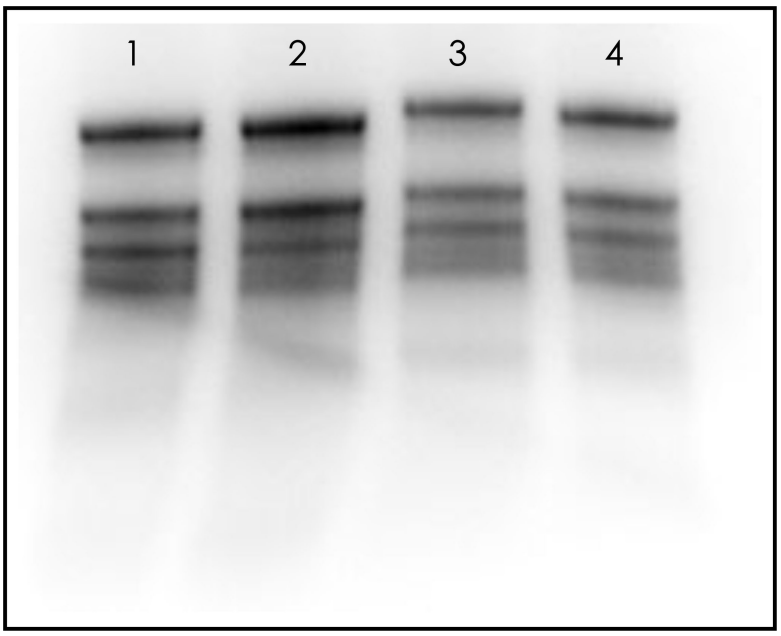


Figure 4. Total RNA extracted from *S. coronatus* leaf tissue using the acid guanidinium thiocyanate-phenol-chloroform extraction method and electrophoresed on a formaldehyde-MOPS RNA denaturing gel. Approximately 2 µg RNA was loaded per lane.

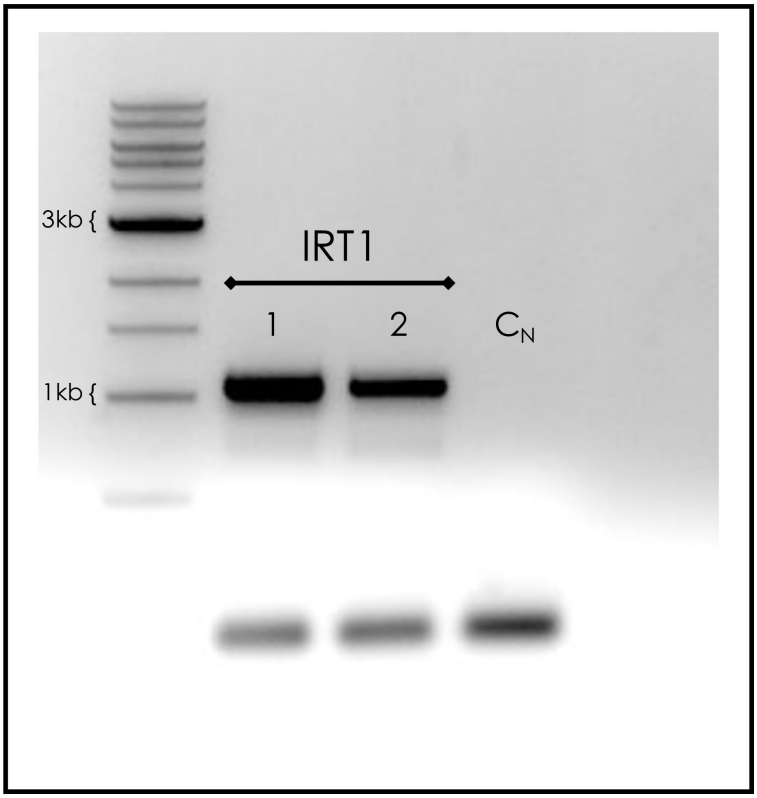


Figure 5. PCR amplification of the full-length *IRT1* transcript from *S. coronatus* cDNA, run out on a 1% (v/v) agarose gel. C_N indicates a negative control (i.e. no-template PCR). Amplicon bands of approximately 1050 bp were expected.

PCR amplification of *MATE*, *VIT* and *COP*, vector preparation and cloning

During the initial process of PCR optimisation, *MATE* and *VIT* were amplified using a standard polymerase (Figure 6). Having refined the optimal PCR parameters for these genes, they were again amplified using a high-fidelity proof-reading polymerase (Figure 7) so as to minimise polymerase-introduced errors. *COP* was amplified with the high fidelity proof-reading polymerase without the need for prior optimisation (Figure 8).

The pENTR1A entry vector was initially digested with *Bam*HI and *Xho*I, but attempts to clone the three GOIs into the *Bam*HI/*Xho*I-digested vector backbone were unsuccessful. Efforts to optimise the ligation reaction were unproductive. Although *Xho*I appeared to cut the pENTR1A entry vector into fragments of the expected size (gel not shown), it seemed unable to efficiently cut the *MATE*, *VIT* and *COP* amplicons. As such, *Xho*I was replaced by *Eco*RI in the digestion process (Figure 9) and new primers (incorporating the *Eco*RI restriction site) were used to amplify the GOIs. Subsequent cloning attempts resulted in the growth of putative positive transformants on selective media.

However, a further challenge was the apparent absence of inserts in the putative positive transformants. Several ligation and transformation attempts were made (using the *Bam*HI/*Eco*RI-digested entry vector) that resulted in colony growth on selective media, but for which colony PCR (see below) revealed no indication of the GOI. This suggests that, although the "sticky ends" resulting from *Bam*HI and *Eco*RI digestion are incompatible, the digested vector was in fact re-circularising. Such a hypothesis was supported when, after dephosphorylation of the digested entry vector, positive transformants resulted. Furthermore, apparent transformations were generated when only vector backbone (i.e. no insert DNA) was subjected to the ligation reaction.

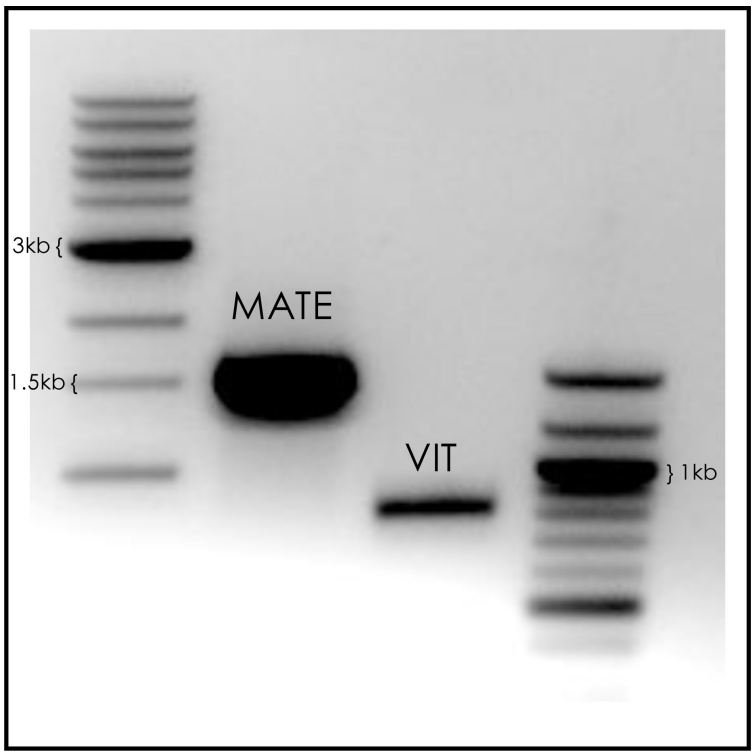


Figure 6. PCR amplification of *MATE* and *VIT* run out on a 1% (w/v) agarose gel. Amplicon bands of approximately 1.5 kb and 800 bp respectively were expected.

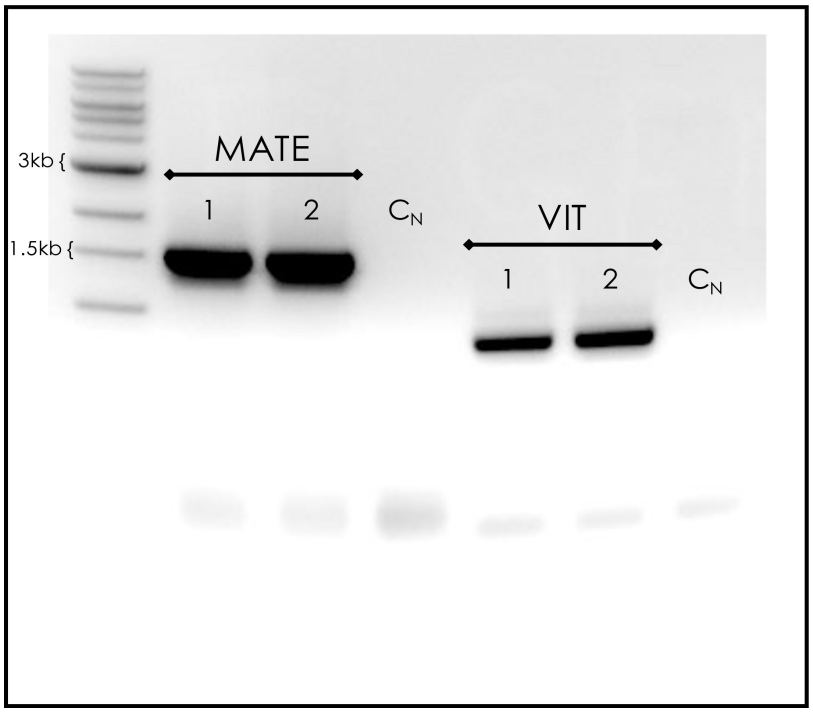


Figure 7. PCR amplification of *MATE* and *VIT* using high-fidelity proof-reading polymerase and run out of a 1% (w/v) agarose gel. C_N indicates a negative control (i.e. no-template PCR). Amplicon bands of approximately 1.5 kb and 800 bp respectively were expected.

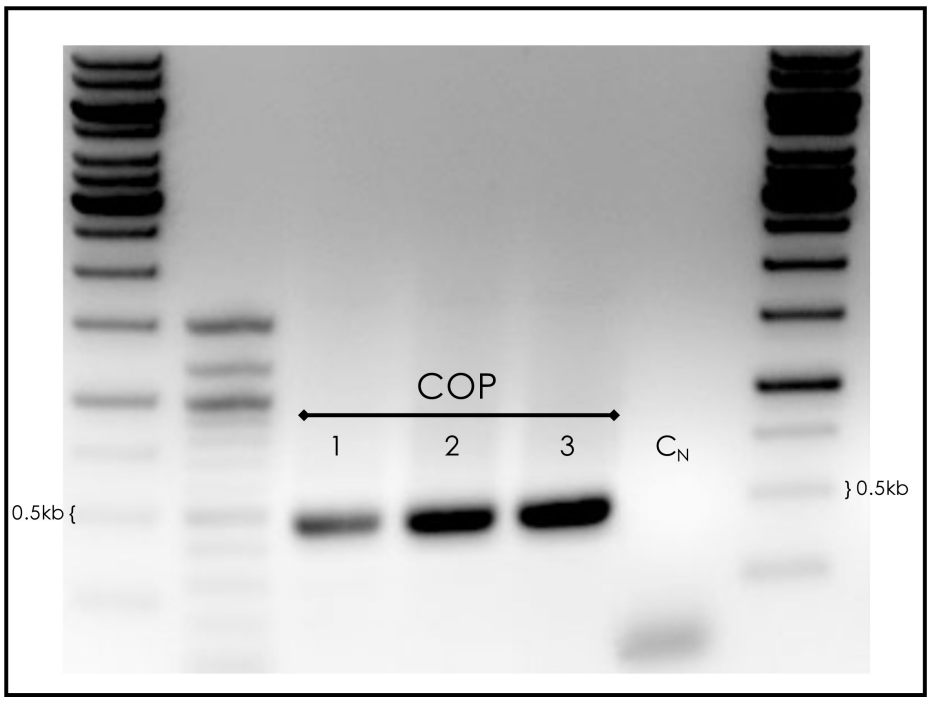


Figure 8. PCR amplification of *COP* using high-fidelity proof-reading polymerase and run out of a 1% (w/v) agarose gel. C_N indicates a negative control (i.e. no-template PCR). Amplicon bands of approximately 500 bp were expected.

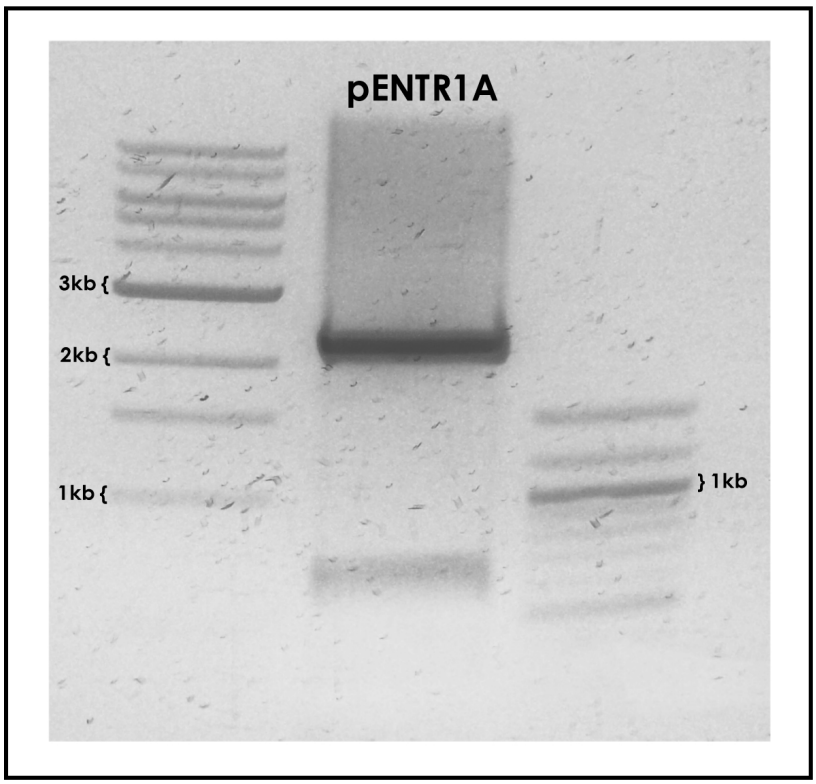


Figure 9. Restriction enzyme digestion of pENTR1A using *Eco*RI and *Bam*HI. A backbone of 2282 bp was expected, along with 5 fragments of 681 bp, 452 bp 250 bp, 77 bp and 12 bp. The smaller fragments are not visible here.

Confirmation of positive transformants

Putative positive transformants were screened for the presence or absence of the relevant GOI using colony PCR with gene-specific primers. Bands of the predicted size were observed following electrophoresis of PCR products for all three GOIs (Figures 10-12). In order to further minimise the possibility of proceeding with false positives, diagnostic double digests were performed on plasmid DNA extracted from overnight cultures of these colonies. In the case of *MATE* and *VIT*, the bands observed correspond exactly with those expected for the backbone and the insert (Figure 13). In the case of *COP*, although the vector backbone was clearly in evidence, the insert was either absent or too faint to visualise on the gel (Figure 14). Given the relatively small expected fragment size (507 bp), the latter is possible. Thus, a PCR (using gene-specific primers) was performed on the plasmid DNA from the putative *COP* transformants and this revealed a band of the expected insert size (Figure 15).

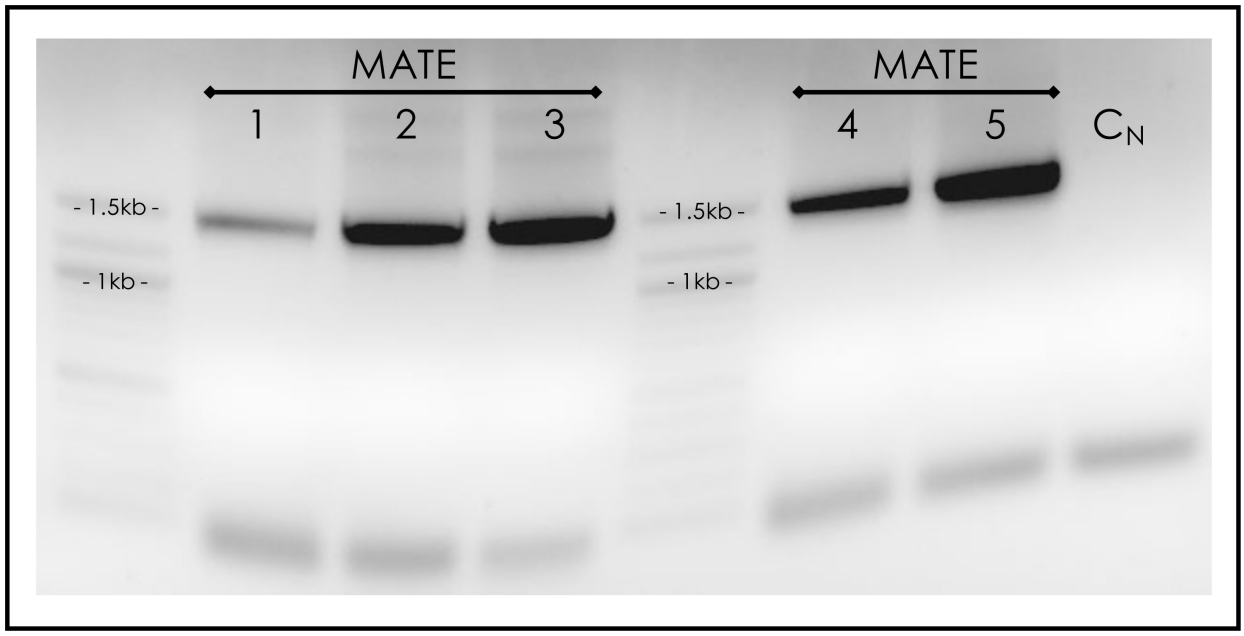


Figure 10. Colony PCR performed on putative positive *E. coli* (DH5α) transformants thought to contain the pENTR1A-*MATE* construct. C_N indicates a negative control (i.e. no-template PCR). Amplicon bands of approximately 1.5 kb were expected.

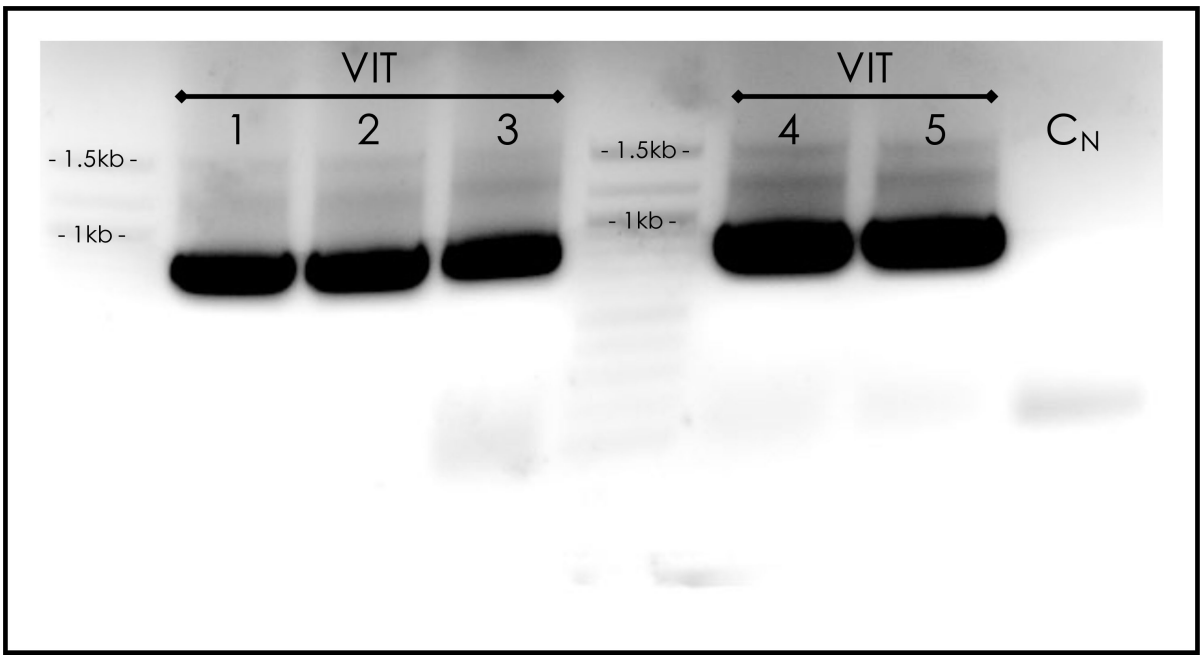


Figure 11. Colony PCR performed on putative positive *E. coli* (DH5α) transformants thought to contain the pENTR1A-*VIT* construct. C_N indicates a negative control (i.e. no-template PCR). Amplicon bands of approximately 800 bp were expected.

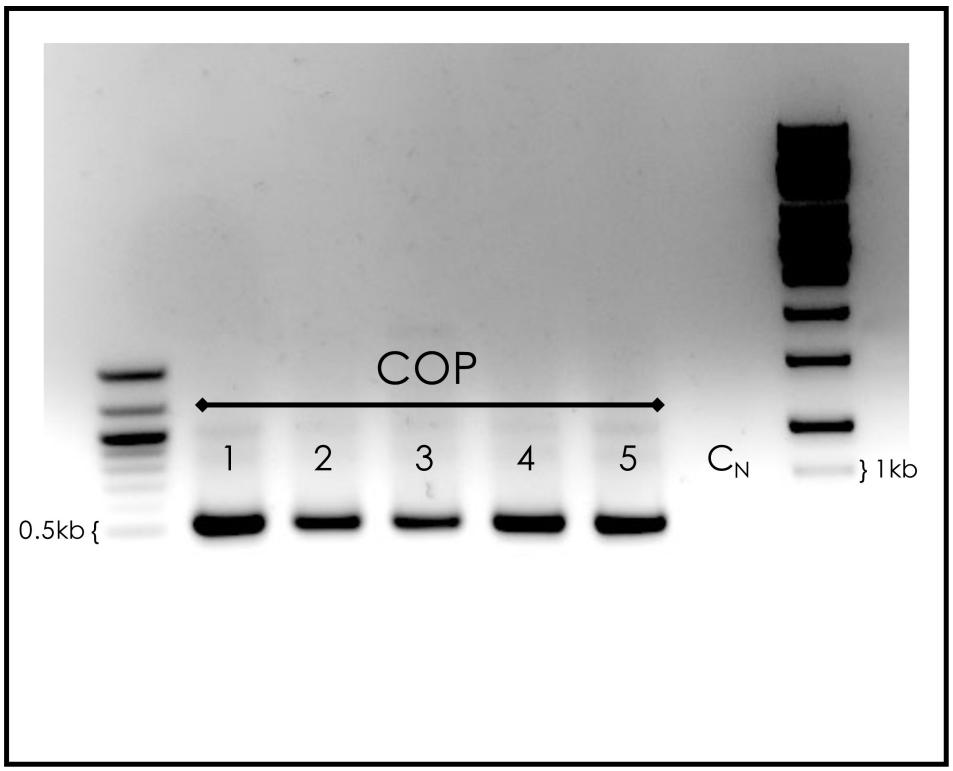


Figure 12. Colony PCR performed on putative positive *E. coli* (DH5α) transformants thought to contain the pENTR1A-*COP* construct. C_N indicates a negative control (i.e. no-template PCR). Amplicon bands of approximately 500 bp were expected.

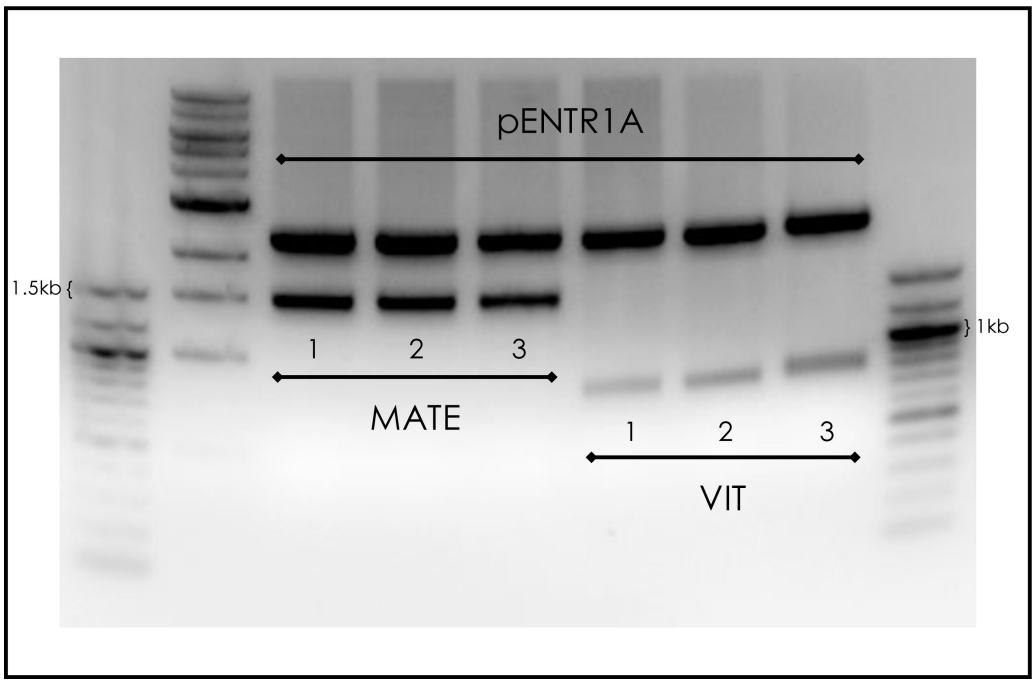


Figure 13. Diagnostic double digestion performed on extracted plasmid from putative positive *E. coli* (DH5α) transformants thought to contain the pENTR1A-*MATE* and pENTR1A-*VIT* constructs. The restriction enzymes *Eco*RI and *Bam*HI were used. A vector backbone of 2282 bp was expected, along with insert sizes of approximately 1.5 kb and 800 bp respectively.

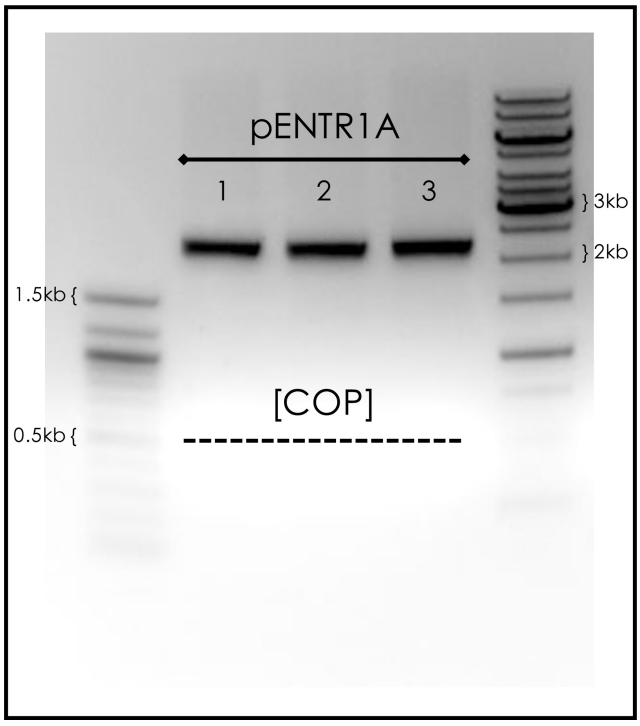


Figure 14. Diagnostic double digestion performed on extracted plasmid from putative positive *E. coli* (DH5α) transformants thought to contain the pENTR1A-*COP* construct. A vector backbone of 2282 bp was expected, along with an insert size of approximately 500 bp.

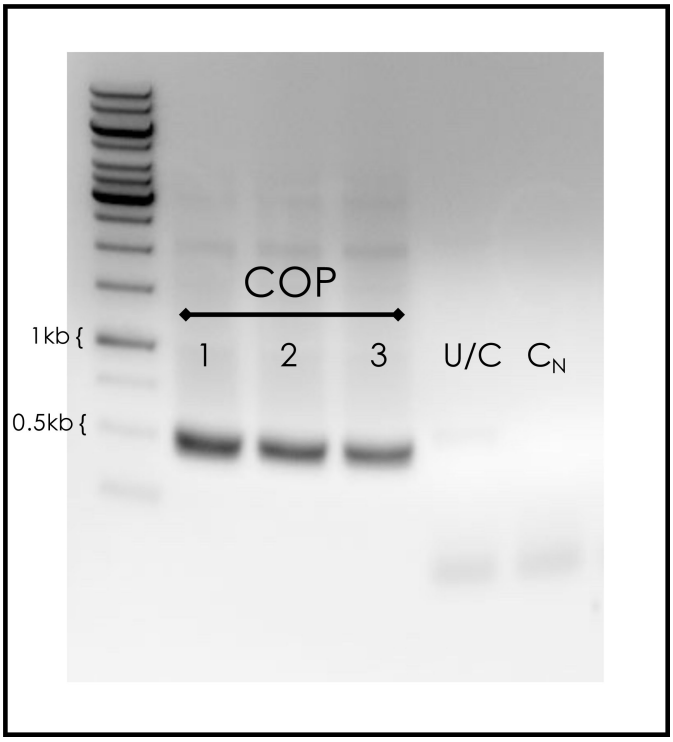


Figure 15. Plasmid PCR performed on extracted plasmid from putative positive *E. coli* (DH5α) transformants thought to contain the pENTR1A-*COP* construct. C_N indicates a negative control (i.e. no-template PCR) and U/C is the empty plasmid (pENTR1A). Amplicon bands of approximately 500 bp were expected.

2. Nucleotide and amino acid sequence analysis

Validation of RNA-Seq derived sequences for putative transport proteins

The nucleotide sequences obtained for the amplified cDNA *ScMATE*, *ScVIT* and *ScCOP* transcripts were aligned against the consensus sequences derived from RNA-Seq *de novo* assembly of the *S. coronatus* transcriptome (Appendix 1). There were several nucleotide substitutions in *ScMATE*, predominantly towards the 3' end of the sequence. In order to determine whether these differences were due to errors introduced by DNA polymerase during the PCR or instead reflect variation between individual *S. coronatus* plants (different plants were used for the RNA-Seq experiment), two further reverse sequences were generated for *ScMATE* (from independent clones) and added to the alignment (Appendix 1). The three reverse sequences aligned perfectly, other than at position 593. Here, one reverse sequence matched the RNA-Seq consensus sequence and the other two reverse sequences were identical. There were no nucleotide substitutions in *ScVIT* (i.e. 100% match to the RNA-Seq nucleotide sequence). There was one nucleotide substitution at position 177 in the *ScCOP* sequence.

In order to determine whether these differences in the nucleotide sequences would influence the amino acid sequence, the cDNA sequences were translated into their corresponding amino acid sequences (Gasteiger et al. 2003), then aligned against the translated RNA-Seq derived sequences. The ScMATE amino acid sequence (Figure 16) revealed four amino acid changes. Two of these substitutions were between functionally similar amino acids (at positions 282 and 405). The ScVIT amino acid sequence was conserved (Figure 17). The nucleotide substitution in *ScCOP* was silent, resulting in a conserved amino acid sequence (Figure 18).

These data therefore validate the consensus sequences generated through RNA-Seq, although the functional significance of the ScMATE amino acid changes remain unclear. A first step in investigating this would be to determine whether the nucleotide changes observed represent the "true" *ScMATE* sequence (i.e. the RNA-Seq derived sequence is erroneous) or are merely a result of natural heterogeneity

at these positions. This could be accomplished by sequencing this gene from a number of individuals across multiple populations of this species.

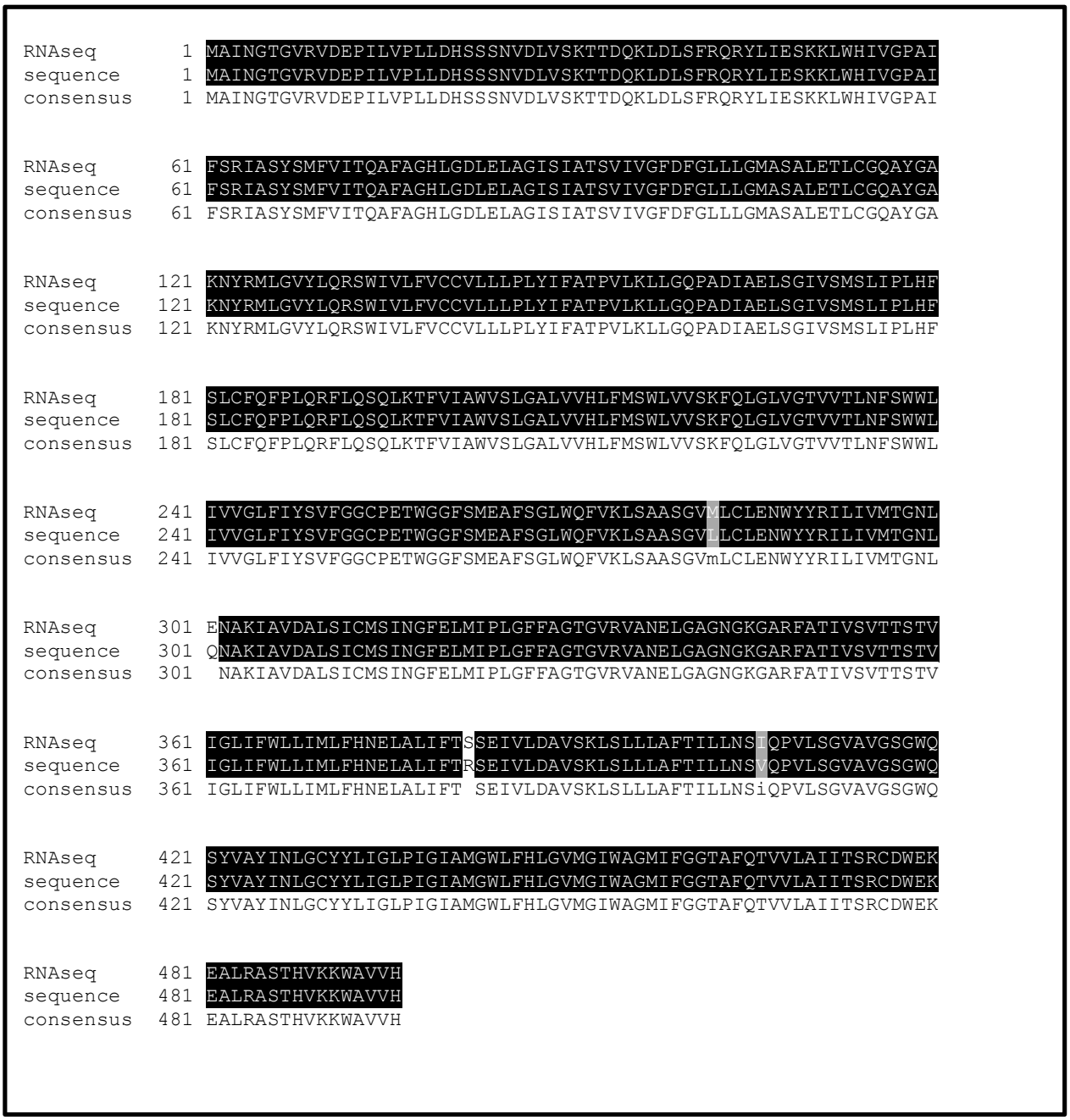


Figure 16. ScMATE amino acid sequence alignment between the RNA-Seq derived sequence ("RNAseq") and RT-PCR derived sequences ("sequence"). Amino acid substitutions are highlighted in grey where substitutions are between amino acids of similar properties and in white where such properties are dissimilar.

```
RNAseq       1 MAGNIGGATITTTPFLTNGAAAEENGSGRERPKEPWKGELVKSIVYAGLDAIVTSFSLIS
sequence     1 MAGNIGGATITTTPFLTNGAAAEENGSGRERPKEPWKGELVKSIVYAGLDAIVTSFSLIS
consensus    1 magniggatitttpfltngaaaeengsgrerpkepwkgelvksivyagldaivtsfslis

RNAseq      61 SISAGRLSSVDVLVLGFANLVADGISMGFGDYVSSNTERDVAAKERLVTEWEVANRRRNQ
sequence    61 SISAGRLSSVDVLVLGFANLVADGISMGFGDYVSSNTERDVAAKERLVTEWEVANRRRNQ
consensus   61 sisagrlSSVDVLVLGFANLVADGISMGFGDYVSSNTERDVAAKERLVTEWEVANRRRNQ

RNAseq     121 EQELLDRYQDLGMNIQDATTVVSIFAKYGDIMVDEKMIHKGTLSPDDGEKPWKKGLITFV
sequence    54 EQELLDRYQDLGMNIQDATTVVSIFAKYGDIMVDEKMIHKGTLSPDDGEKPWKKGLITFV
consensus  121 EQELLDRYQDLGMNIQDATTVVSIFAKYGDIMVDEKMIHKGTLSPDDGEKpwkkglitfv

RNAseq     181 AFLVFGSAPILAFIILIPFTHNDTHKFIGACILSALALAALGIAKAKIAGQNHMLSAGGT
sequence   114 AFLVFGSAPILAFIILIPFTHNDTHKFIGACILSALALAALGIAKAKIAGQNHMLSAGGT
consensus  181 aflvfgsapilafiilipfthndthkfigacilsalalaalgiakakiagqnhmlsaggt

RNAseq     241 LFNGALAGFAAYAIGWVLRDVAGLED
sequence   174 LFNGALAGFAAYAIGWVLRDVAGLED
consensus  241 lfngalagfaayaigwvlrdvagled
```

Figure 17. ScVIT amino acid sequence alignment between the RNA-Seq derived sequence ("RNAseq") and the RT-PCR derived sequences ("sequence").

```
RNAseq       1 MMHMTFCWGTNVTLLIDSWKTDSWFSYSLALIICFIFSAFYQFMEDIRLRFKLLSSSVTA
forward      1 MMHMTFCWGTNVTLLIDSWKTDSWFSYSLALIICFIFSAFYQFMEDIRLRFKLLSSSVTA
consensus    1 MMHMTFCWGTNVTLLIDSWKTDSWFSYSLALIICFIFSAFYQFMEDIRLRFKLLSSSVTA

RNAseq      61 VGAVENAPLIYNKFFSGGRRARFAGSLLFGINSGLNYFLMLAVMSFNVGVFVVIVAGLAV
forward     61 VGAVENAPLIYNKFFSGGRRARFAGSLLFGINSGLNYFLMLAVMSFNVGVFVVIVAGLAV
consensus   61 VGAVENAPLIYNKFFSGGRRARFAGSLLFGINSGLNYFLMLAVMSFNVGVFVVIVAGLAV

RNAseq     121 GYWLFRSADDEQITLL
forward    121 GYWLFRSADDEQITLL
consensus  121 GYWLFRSADDEQITLL
```

Figure 18. ScCOP amino acid sequence alignment between the RNA-Seq derived sequence ("RNAseq") and the RT-PCR derived sequences ("sequence").

Phylogenetic analyses

The ScMATE, ScVIT and ScCOP sequences were annotated based on amino acid sequence similarity during the analysis of the RNA-Seq data. In order to examine their evolutionary relationships with other plant proteins, phylogenetic analysis was performed. The amino acid sequences were used in a BLASTP search against the NCBI Viridiplantae database. Seven homologous amino acid sequences were selected from different plant species in the top 10-15 BLAST hits. Where these results included a sequence in *N. caerulescens*, it was automatically incorporated into this initial selection. Where it did not, the BLASTP search set was limited by organism to *Noccaea*. This was done in order to incorporate a second hyperaccumulator in all three phylogenies. Similarly constrained searches were performed against both *A. thaliana* and *Zea mays*. The best match from each of these constrained searches was added to the selection, making a total of nine amino acid sequences for each ScMATE, ScVIT and ScCOP.

Taxonomic data, accession numbers and BLAST scores are presented in Table 3. The MATE amino acid sequences shared high sequence identity across all taxa; the mean sequence identity was 74.8%. This was less pronounced with VIT (59.2%), although if *Arabidosis*, *Noccaea* and *Zea* were excluded, the mean sequence identity was 75.5%. Sequence homology was less pronounced with COP, which had a mean sequence identity of 54.3% across all taxa. Given that amino acid identities as low as 32% have been considered as indicative of protein relatedness (Shitan et al. 2014), these data strongly support the notion that ScMATE, ScVIT and ScCOP are homologs of other, known metal transport proteins.

These amino acid sequences were then aligned using MEGA7 (alignments not included here) and used to generate the neighbour-joining phylogenies shown in Figures 19-21. Sequence homologies appear to mirror known family-level evolutionary relationships in all phylogenies with one notable exception: *A. thaliana* and *N. caerulescens* form a well-supported clade for VIT and COP but not MATE. As the only brassicids in the analysis, one would expect that their close evolutionary

relatedness would be reflected in their sequence similarity, although a lack of comprehensive genomic data for *N. caerulescens* limits speculation in this regard. By way of contrast, in the VIT phylogeny, all the solanacids form a well-supported clade. It is interesting to note, however, that the instance in which *N. caerulescens* does not group with its fellow brassicid, it is grouping with *S. coronatus* – the only other hyperaccumulator in the analysis.

Table 3. Amino acid homologs for ScMATE, ScVIT and ScCOP as optained from a BLASTp search. Percentage coverage, shared identity and E values as indicated.

Species	Family	Accession	Coverage	Identity	E value
MATE					
Arabidopsis thaliana	Brassicaceae	NP_201341.1	95%	72%	0
Citrus sinensis	Rutaceae	XP_006466280.1	97%	77%	0
Gossypium arboreum	Malvaceae	XP_017627837.1	96%	78%	0
Gossypium hirsutum	Malvaceae	XP_016679591.1	96%	78%	0
Malus domestica	Rosaceae	XP_008383420.1	99%	76%	0
Morus notabilis	Moraceae	XP_010106306.1	96%	76%	0
Noccaea caerulescens	Brassicaceae	JAU20520.1	99%	69%	0
Sesamum indicum	Pedaliaceae	XP_011097507.1	97%	77%	0
Zea mays	Poaceae	ONL94427.1	89%	71%	0
VIT					
Arabidopsis thaliana	Brassicaceae	BT_028987.1	87%	26%	4e-17
Capsicum annuum	Solanaceae	XP_016574843.1	98%	74%	4e-131
Nicotiana attenuata	Solanaceae	XP_019242015.1	95%	76%	2e-132
Nicotiana tabacum	Solanaceae	XP_016473854.1	95%	76%	4e-133
Nicotiana tomentosiformis	Solanaceae	XP_009601264.1	95%	75%	4e-132
Noccaea caerulescens	Brassicaceae	JAU39496.1	78%	25%	1e-05
Solanum lycopersicum	Solanaceae	XP_004229857.1	88%	78%	2e-129
Solanum tuberosum	Solanaceae	XP_006339489.1	95%	74%	7e-130
Zea mays	Poaceae	XP_008663753.1	31%	29%	0.007
COP					
Arabidopsis thaliana	Brassicaceae	NP_197565.1	94%	56%	2e-42
Coffea canephora	Rubiaceae	CDP13483.1	95%	54%	3e-44
Erythranthe guttata	Phrymaceae	XP_012857960.1	97%	56%	1e-46
Ipomoea nil	Convolvulaceae	XP_019188867.1	100%	59%	3e-49
Nicotiana attenuata	Solanaceae	XP_019238372.1	99%	55%	1e-44
Nicotiana sylvestris	Solanaceae	XP_009800304.1	99%	55%	1e-44
Noccaea caerulescens	Brassicaceae	JAU53374.1	94%	56%	3e-39
Solanum lycopersicum	Solanaceae	XP_004232609.1	99%	56%	1e-45
Zea mays	Poaceae	ONM22193.1	100%	42%	7e-34

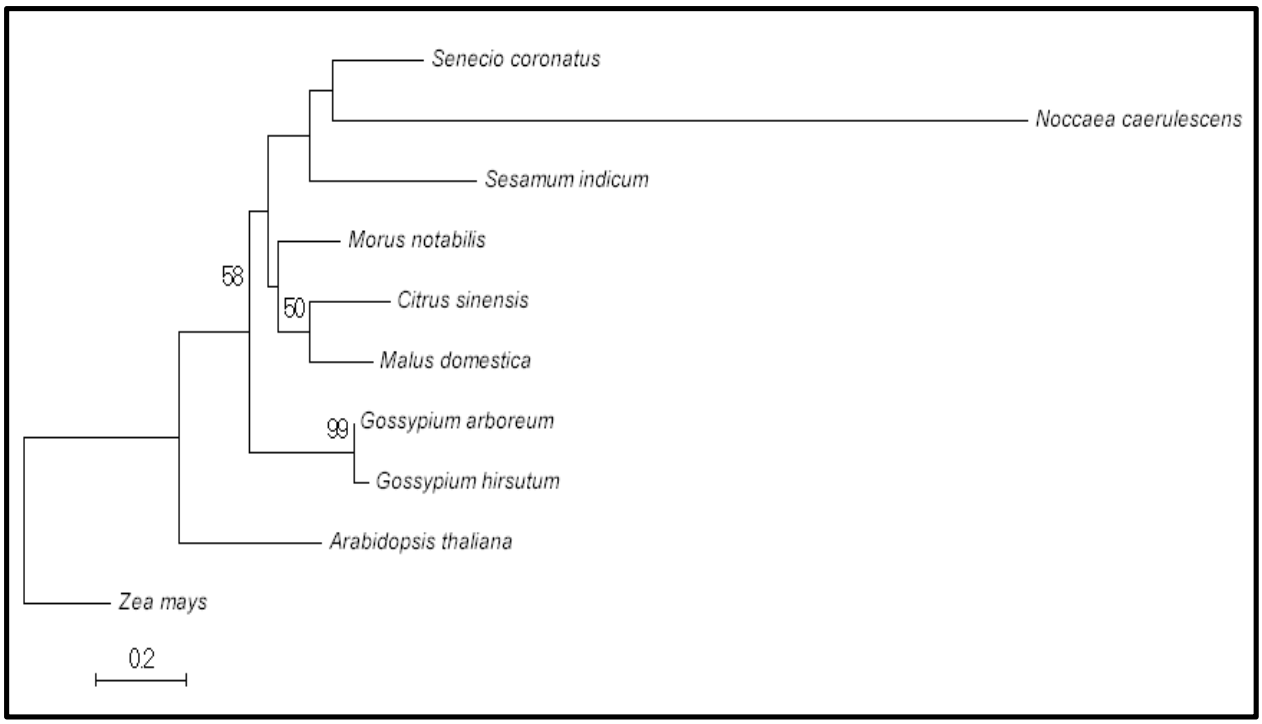


Figure 19. Neighbour-joining tree representing phylogenetic relationships in amino acid sequences homologous to *S. coronatus* MATE. The evolutionary history was inferred using the Neighbour-Joining method (Saitou & Nei 1987). The optimal tree with the sum of branch length = 4.13605079 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (10000 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl & Pauling 1965). The analysis involved 10 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 40 positions in the final dataset. The tree was rooted on *Zea mays*. Evolutionary analyses were conducted in MEGA7 (Kumar et al. 2016).

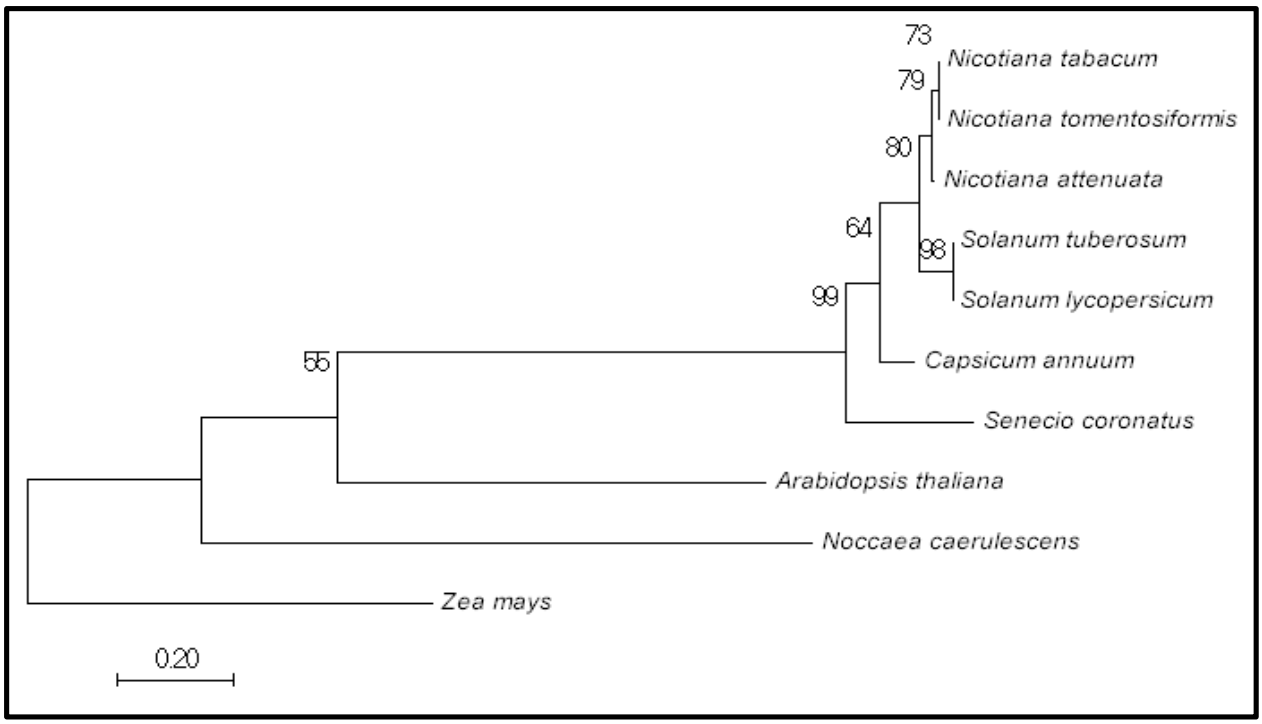


Figure 20. Neighbour-joining tree representing phylogenetic relationships in amino acid sequences homologous to *S. coronatus* VIT. The evolutionary history was inferred using the Neighbour-Joining method (Saitou & Nei 1987). The optimal tree with the sum of branch length = 4.37880026 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (10000 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl & Pauling 1965). The analysis involved 10 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 46 positions in the final dataset. The tree was rooted on *Zea mays*. Evolutionary analyses were conducted in MEGA7 (Kumar et al. 2016).

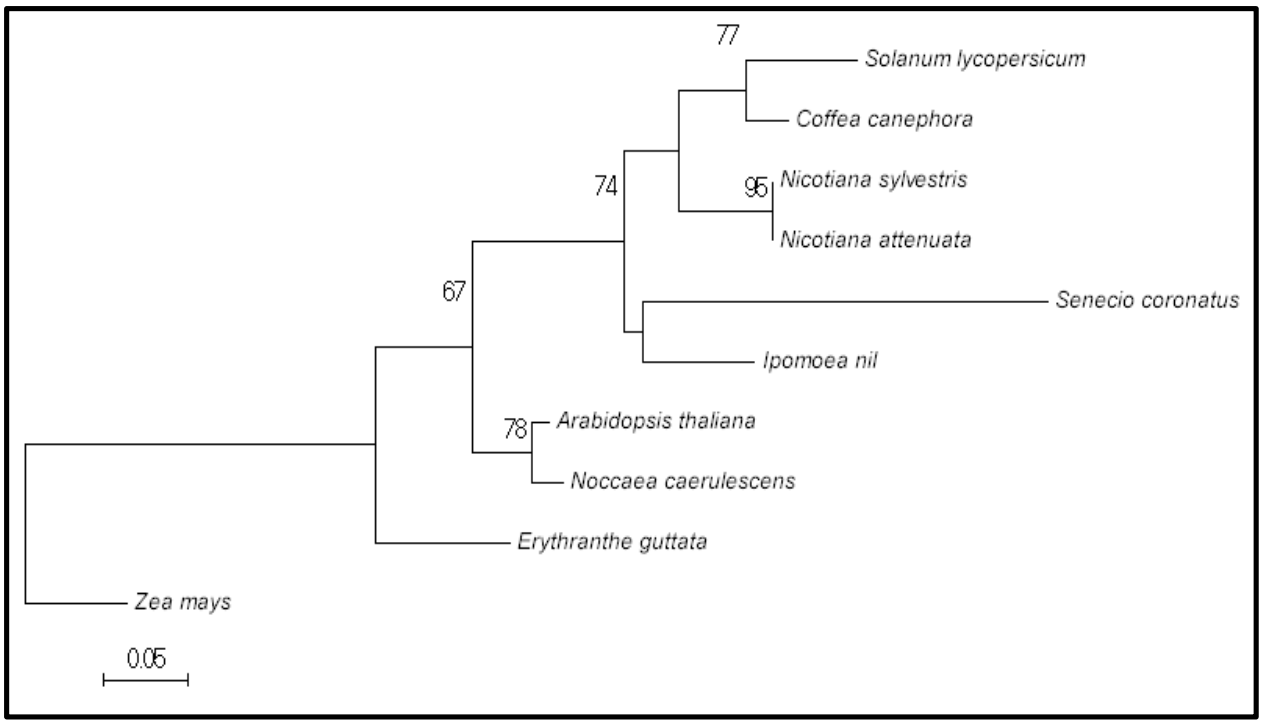


Figure 21. Neighbour-joining tree representing phylogenetic relationships in amino acid sequences homologous to *S. coronatus* COP. The evolutionary history was inferred using the Neighbour-Joining method (Saitou & Nei 1987). The optimal tree with the sum of branch length = 1.10022486 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (10000 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl & Pauling 1965). The analysis involved 10 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 34 positions in the final dataset. The tree was rooted on *Zea mays*. Evolutionary analyses were conducted in MEGA7 (Kumar et al. 2016).

Protein topology predictions

The difficulty inherent in experimentally mapping transmembrane protein topology is widely-acknowledged (Käll et al. 2004). This has resulted in the emergence of a plethora of predictive software tools of varying degrees of sophistication. Having established above that ScMATE, ScVIT and ScCOP share sequence similarities with the homologous proteins in other species, two commonly used protein topology prediction platforms, Phobius (Krogh et al. 2001) and the TMHMM Server (Käll et al. 2004), were used to determine whether the *S. coronatus* proteins contain putative transmembrane domains ("TMD") and whether these match those predicted for *Arabidopsis*. A survey of the current literature reveals that MATE proteins are typically predicted to have twelve TMDs (Braibant et al. 2002; Brown et al. 1999; Maron et al. 2010; Shitan et al. 2014), VIT proteins five TMDs (Slavic et al. 2016; Zhang et al. 2012) and COP/T proteins three TMDs (Puig & Thiele 2002).

The Phobius predictions are shown in Figures 22-24 and the TMHMM predictions in Appendix 2. The Phobius predictions for ScMATE and AtMATE were consistent with those predicted in the literature (i.e. 12 transmembrane domains). The topology patterns mirrored those predicted by TMHMM, although TMHMM predicted only 10 transmembrane domains for both ScMATE and AtMATE. This is presumably due to the application of a stricter probability threshold. Similarly, both ScVIT and AtVIT Phobius predictions matched those in the literature (5 transmembrane domains). TMHMM predicted 3 TMDs for ScVIT and 5 for AtVIT, despite having very similar topographies. Both prediction platforms predicted 3 transmembrane domains for ScCOP and AtCOP, consistent with the literature.

The consistency between the predicted number of TMDs in ScMATE, ScVIT and ScCOP and their *Arabidopsis* homologs (which matches predictions made in the literature) – and the close topological similarities between them – further supports the idea that MATE, VIT and COP may act as metal transport proteins in *S. coronatus*.

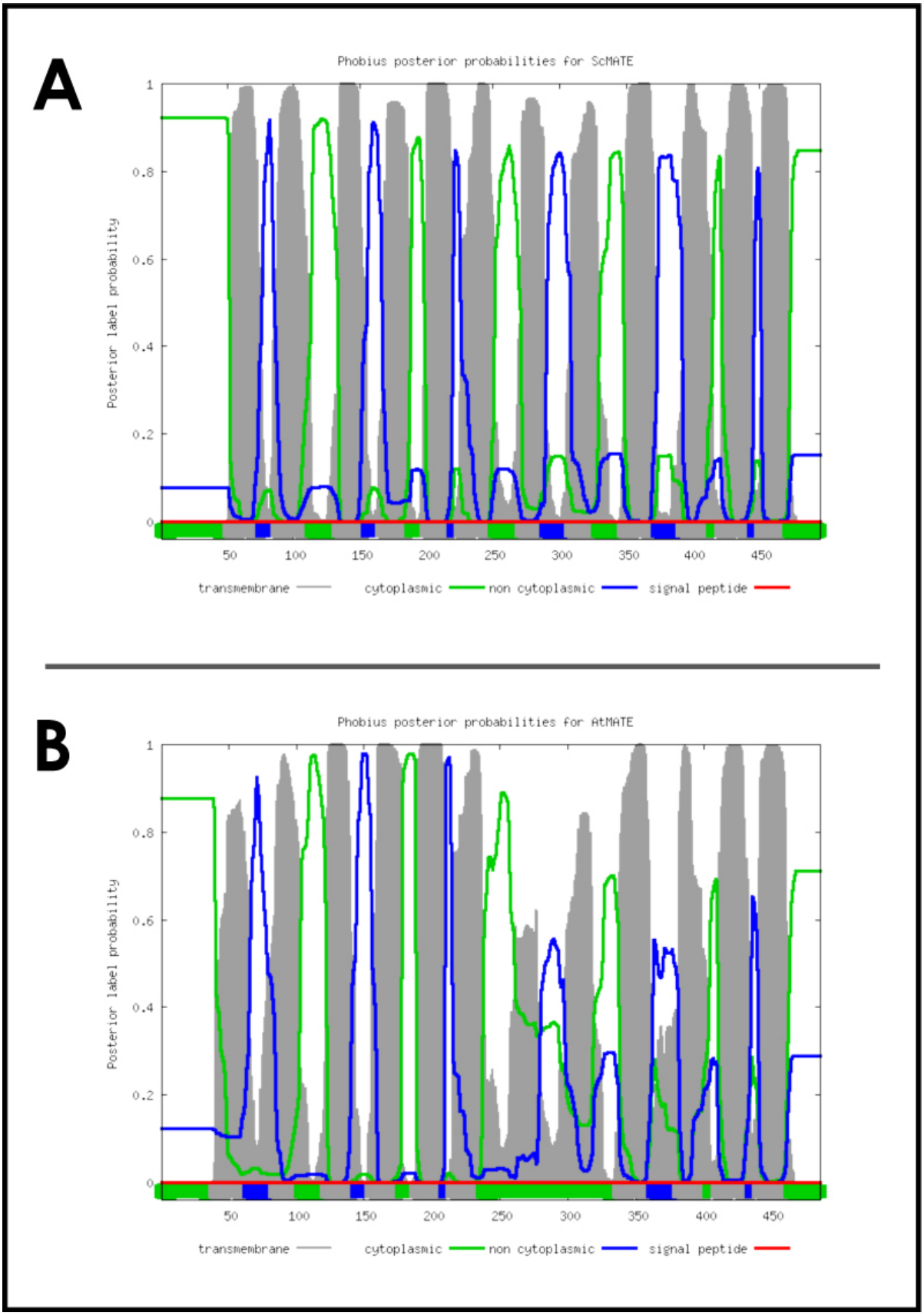


Figure 22. Transmembrane domain predictions for ScMATE (**A**) and AtMATE (**B**) generated using Phobius. The total number of TMDs predicted is twelve.

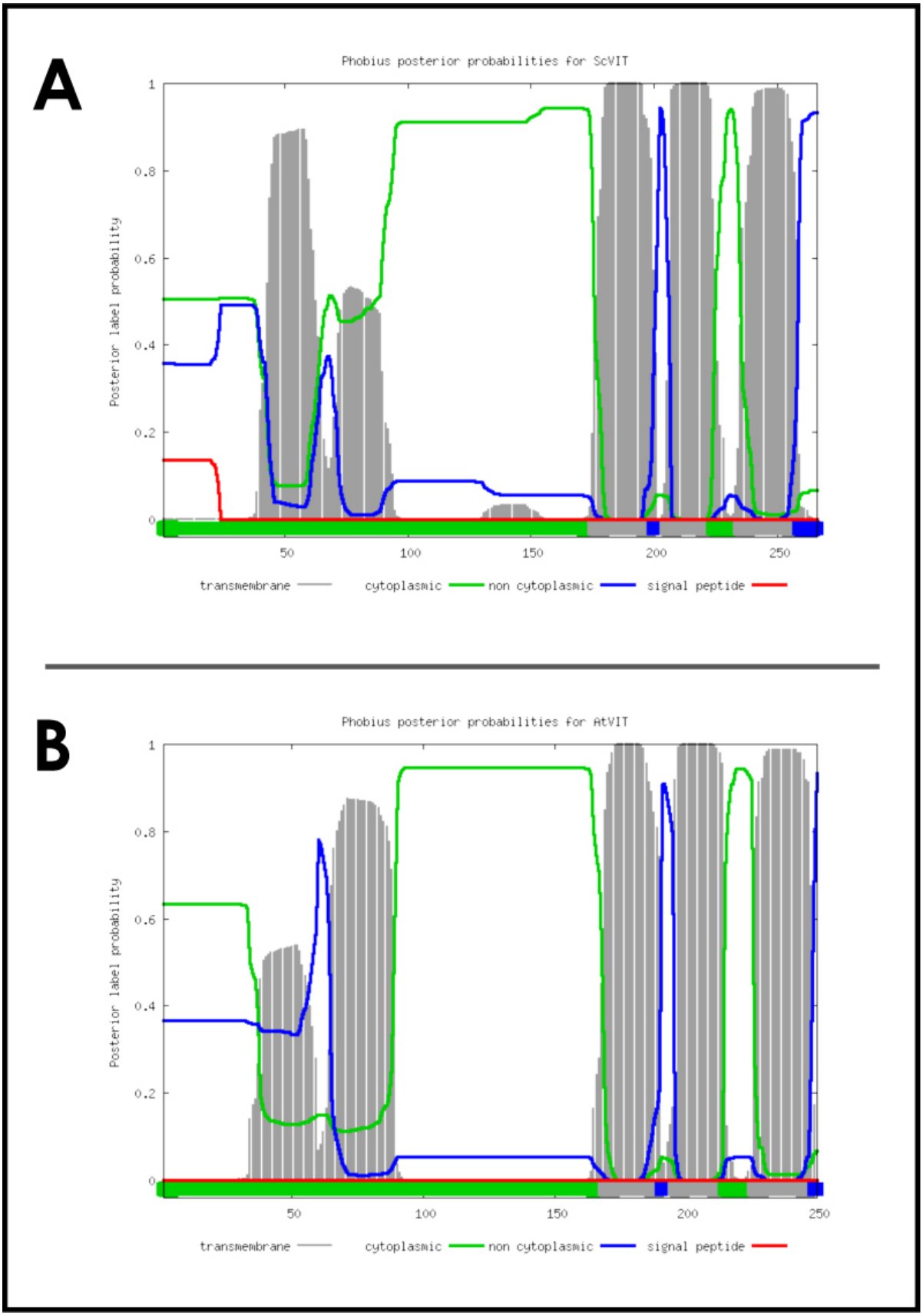


Figure 23. Transmembrane domain predictions for ScVIT (**A**) and AtVIT (**B**) generated using Phobius. The total number of TMDs predicted is five.

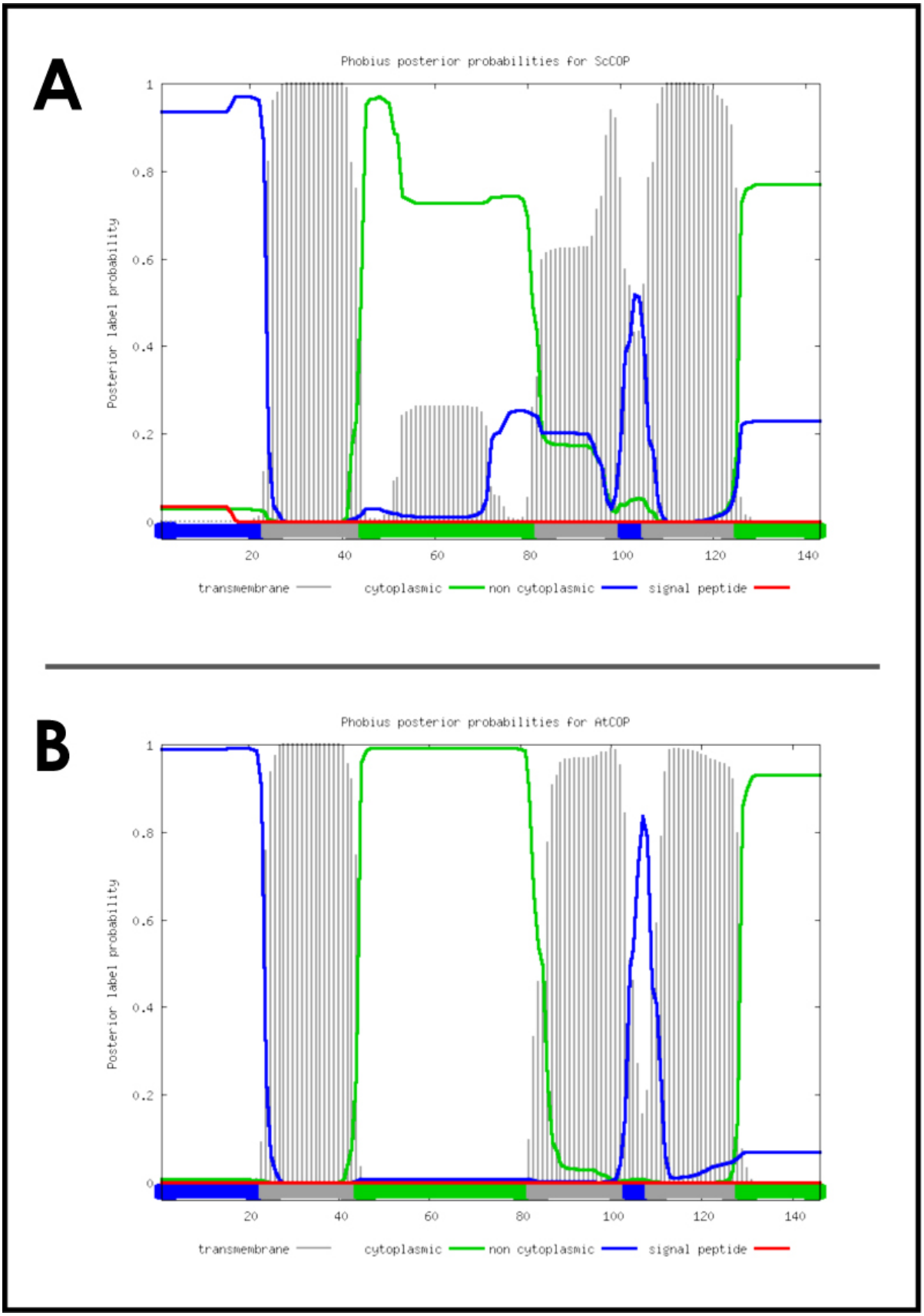


Figure 24. Transmembrane domain predictions for ScCOP (**A**) and AtCOP (**B**) generated using Phobius. The total number of TMDs predicted is three.

Protein localisation predictions

The function of a protein is necessarily dependent on its location within a cell (Yu et al. 2008). Although subcellular protein localisation is amenable to experimental investigation, the interpretation of such data can be problematic depending on the specific techniques used (Hooper et al. 2014). As such, the accurate prediction of subcellular protein localisation has value in the characterisation and functional annotation of proteins (Chang et al. 2013). Unfortunately, however, localisation predictors are notoriously inconsistent (Hooper et al. 2014). This has led to the development of software capable of aggregating predictions from various software platforms along with data from mass spectrometry, GFP localisation, indirect protein-protein interaction and co-expression experiments, producing a rigorously supported "consensus call" (Hooper et al. 2014; Tanz et al. 2013). Currently this software is applicable only to the *Arabidopsis* proteome and was used here to make predictions about the Arabidopsis homologs of ScMATE, ScVIT and ScCOP. This allowed for a more nuanced evaluation of the less sophisticated platforms, which can be used to make predictions about *S. coronatus* protein localisations.

Figures 25, 26 and 27 contain the key data produced by SUBAcon for AtMATE, AtVIT and AtCOP. What is immediately striking is the inconsistency in the predictions generated by the various prediction platforms (see 25-27 **B**). Nonetheless, the consensus predictions for MATE and COP indicate localisation in the plasma membrane, whereas VIT is predicted to target to the vacuole. MATE and COP proteins have been shown to localise to the plant plasma membrane (Magalhaes et al. 2007; Perea-García et al. 2013). VIT has been experimentally shown to localise in the tonoplast (Kim et al. 2006; Zhang et al. 2012; Zhu et al. 2016).

Despite the marked variation in localisation predictions, three common prediction platforms were used to generate localisation predictions for ScMATE, ScVIT and NcCOP and their homologs in *A. thaliana* and *N. caerulescens* (Table 4). The top two predictions for MATE, VIT and COP were recorded for *S. coronatus*, *A. thaliana* and *N. caerulescens*. However, it was only for MATE that there was unanimity in the

predictions across all three platforms (Chang et al. 2013; Horton et al. 2007; Yu et al. 2008). This prediction, that MATE localises in the plasma membrane, does accord with the consensus prediction made for *Arabidopsis* by SUBAcon. However, there was no such unanimity for VIT and COP. Only one of the three platforms predicted that VIT (across all taxa) would localise to the tonoplast, as is predicted for *Arabidopsis* by SUBAcon and supported by experimental data (Zhu et al. 2016). There was a similar lack of consensus for COP.

These data support the notion that amino acid sequences are themselves inadequate for the consistent and accurate prediction of subcellular protein localisation. Experimental data in this regard is, therefore, essential and is presented in the following sections.

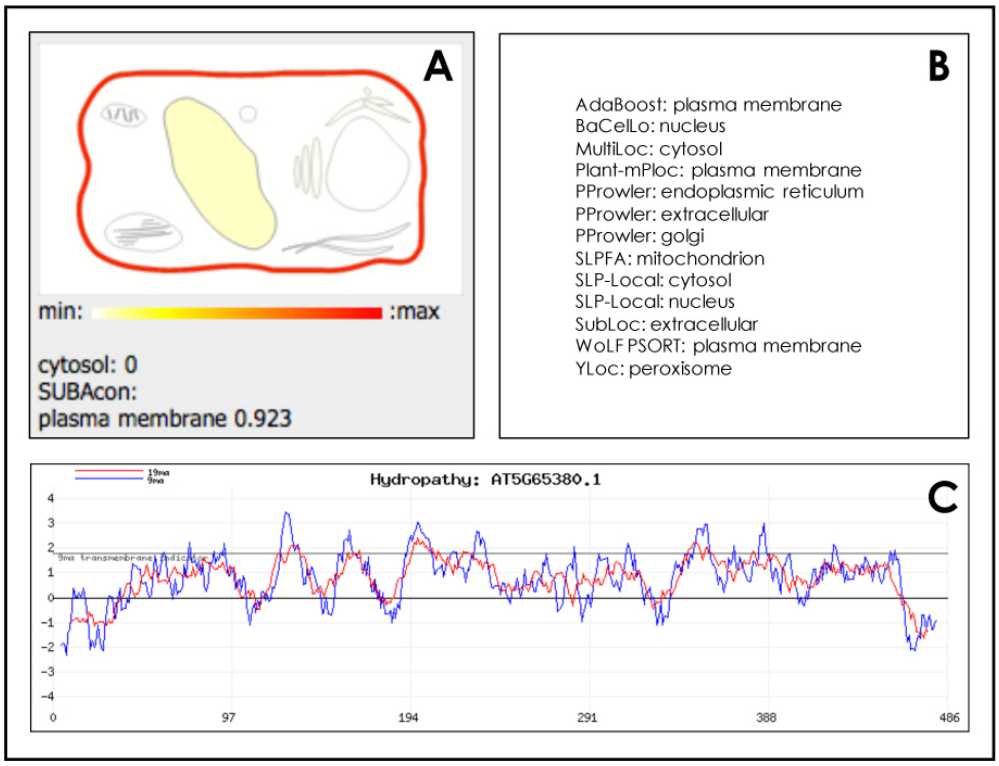


Figure 25. Key protein localisation (**A** and **B**) and topography (**C**) data generated by SUBAcon for the *Arabidopis* homolog (AT5G65380.1) of ScMATE. The prediction shown in **A** represents the consensus prediction for this protein.

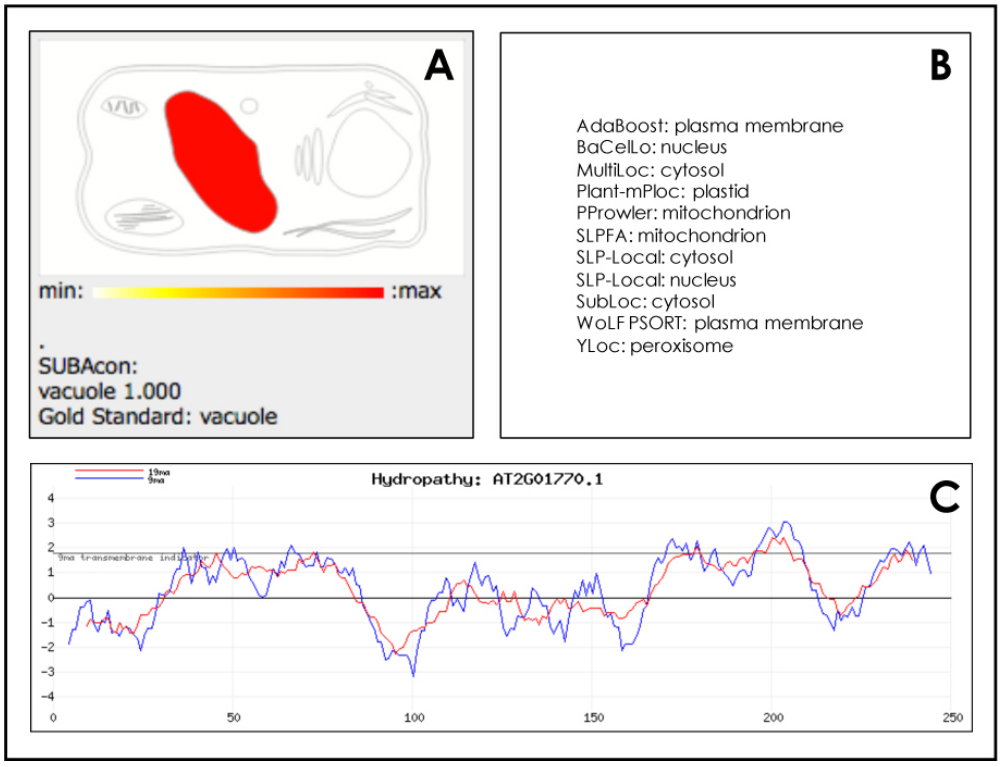


Figure 26. Key protein localisation (**A** and **B**) and topography (**C**) data generated by SUBAcon for the *Arabidopis* homolog (AT4G27860.1) of ScVIT. The prediction shown in **A** represents the consensus prediction for this protein.

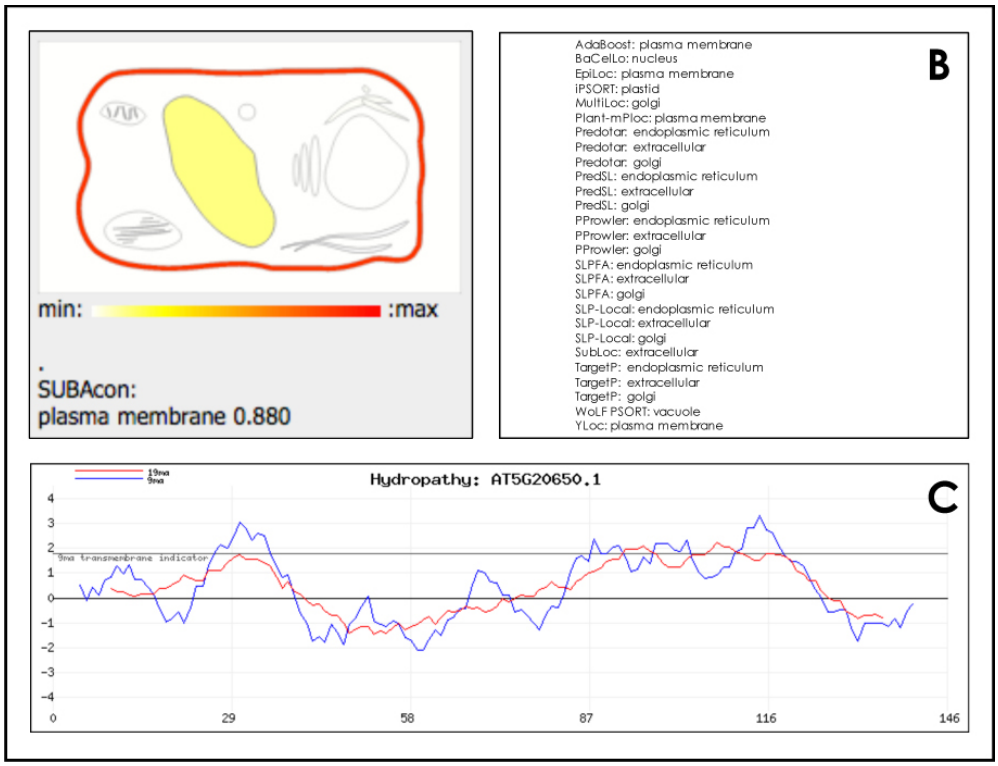


Figure 27. Key protein localisation (**A** and **B**) and topography (**C**) data generated by SUBAcon for the *Arabidopis* homolog (AT5G20650.1) of ScCOP. The prediction shown in **A** represents the consensus prediction for this protein.

Table 4. Subcellular protein localisation predictions and scores for ScMATE, ScVIT and ScCOP and their homologs in *A. thaliana* and *N. caerulescens* generated across three prediction software platforms.

	WoLF PSORT		**Cello**		**EuLoc**	
ScMATE	Plasma membrane	7	Plasma membrane	4.985	Plasma membrane	-
	Tonoplast	5	Peroxisome	0.005	Vacuole	-
AtMATE	Plasma membrane	12	Plasma membrane	4.980	Plasma membrane	-
	Tonoplast	1	Extracellular	0.007	Vacuole	-
NcMATE	Plasma membrane	7	Plasma membrane	4.986	Plasma membrane	-
	Endo. reticulum[53]	4	Nuclear	0.003	Vacuole	-
ScVIT	Tonoplast	9	Plasma membrane	3.496	Plasma membrane	-
	Plasma membrane	2	Chloroplast	0.690	-	-
AtVIT	Plasma membrane	10	Nuclear	3.000	Nuclear	-
	Nuclear	2	Plasma membrane	1.225	-	-
NcVIT	Plasma membrane	10	Nuclear	3.416	Nuclear	-
	Nuclear	2	Cytoplasmic	0.646	-	-
ScCOP	Tonoplast	8	Plasma membrane	4.914	Plasma membrane	-
	Extracellular	3	Chloroplast	0.029	-	-
AtCOP	Tonoplast	8	Plasma membrane	3.954	Plasma membrane	-
	Extracellular	3	Extracellular	0.454	-	-
NcCOP	Endo. reticulum	4	Plasma membrane	4.123	Plasma membrane	-
	Chloroplast	2	Chloroplast	0.221	-	-

[53] Endoplasmic reticulum.

3. Nickel sensitivity of yeast cells expressing *S. coronatus* proteins

MATE, VIT and COP proteins display broad substrate specificity across a wide range of taxa, although are frequently associated with the transport of divalent metal cations (Liu et al. 2016; Magalhaes et al. 2007; Perea-García et al. 2013; Slavic et al. 2016). However, surprisingly little is known about the proteins responsible for nickel transport in plant cells (Merlot et al. 2014). In *S. coronatus*, *MATE*, *VIT* and *COP* have been shown to be constitutively upregulated in hyperaccumulating populations (Meier et al. in review) and it is thus hypothesised that they may be involved in nickel transport.

The interpretation of results from yeast heterologous expression studies can in itself present difficulties. For example, in at least one study involving the heterologous expression of NRAMP4 transport proteins from two *Noccaea* species in yeast, opposing results were obtained: heterologous expression of NRAMP4 from *N. cochleariforme* resulted in increased nickel sensitivity and was associated with greater nickel accumulation (Mizuno et al. 2005). Conversely, decreased sensitivity and reduced nickel uptake was reported in yeast expressing NRAMP4 from *N. caerulescens* (Wei et al. 2009). Furthermore, it has recently been demonstrated that an IREG protein from the nickel hyperaccumulator *Psychotria gabriellae* functioned as a nickel exporter when expressed in yeast, but when expressed in *Arabidopsis* lead to increased vascular nickel concentrations (Merlot et al. 2014). As such, it may not always be possible to infer the direction of transport in *planta* from yeast expression studies, although it is possible to determine a protein's capacity to transport a particular metal ion.

In order to evaluate the capacity of ScMATE, ScVIT and ScCOP to transport nickel, these proteins were heterologously expressed in yeast and any changes in yeast nickel-sensitivity assessed. A yeast double mutant deficient in both high- and low-affinity zinc transporters (*zrt1zrt2*) was used, as the increased sensitivity to metals compared with wild-type yeast serves to magnify changes caused by transgene expression (Merlot et al. 2014).

Expression vector preparation and yeast transformations

Having successfully amplified and cloned *MATE*, *VIT* and COP into the entry vector pENTR1A (see Sections 1 and 2), the Gateway cassette containing these inserts was cloned into the yeast expression vector pDR195 immediately downstream of the *pPMA1* promoter. The recombinant plasmids (pDR195-*ScMATE*, pDR195-*ScVIT* and pDR195-*ScCOP*) were transformed into *E. coli* for propagation and of the presence of inserts confirmed by colony PCR with gene-specific primers (Figure 28).

The extracted plasmids were then transformed into *zrt1zrt2* yeast, where the presence of the GOI was again confirmed by colony PCR with gene-specific primers (Figures 29-31). The presence of the empty pDR195 plasmid in transformed yeast was also confirmed through colony PCR with vector-specific primers (Figure 32). The pDR195-transformed (i.e. empty vector) yeast were used as a negative control in the nickel sensitivity assay, allowing for the establishment of a nickel toxicity baseline. Conversely, yeast transformed with *AtIREG2* were used a positive control, having previously been shown to increase nickel resistance comparative to yeast containing the empty pDR195 vector (Merlot et al. 2014).

It might be noted that yeast colony PCRs were initially unreliable, seldom producing PCR products from digested putative positive transformant yeast colonies. However, it was found that increasing the concentration of sodium hydroxide solution from 20 mM to 40 mM in the initial digestion step substantially improved the success rate of these PCRs.

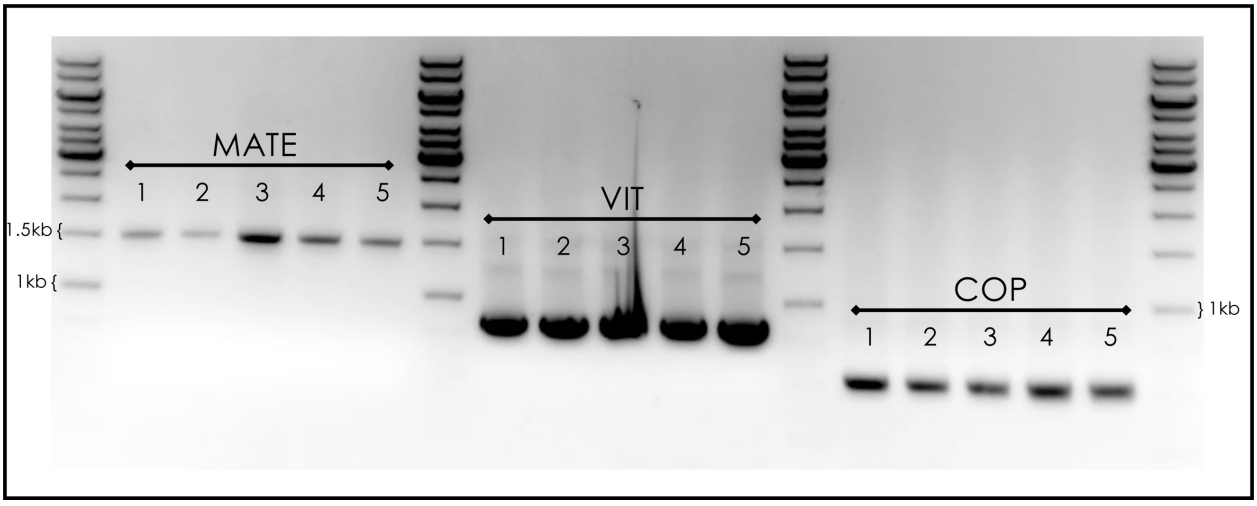


Figure 28. Colony PCR performed on putative positive *E.coli* (DH5α) transformants thought to contain the pDR195-*MATE*, pDR195-*VIT* and pDR195-*COP* constructs. Amplicon bands of approximately 1.5 kb, 800 bp and 500 bp respectively were expected.

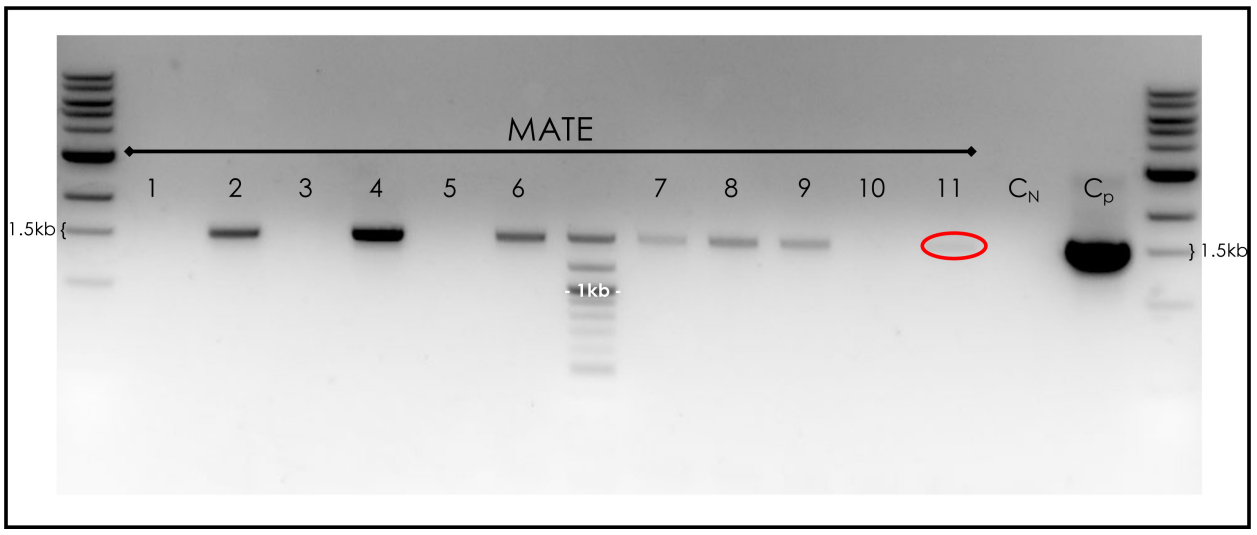


Figure 29. Yeast colony PCR performed on putative positive *S. cerevisiae* (*zrt1zrt2*) transformants thought to contain the pDR195-*MATE* construct. C_N indicates a negative control (i.e. no-template PCR). C_p indicates a positive control (pDR195-*MATE* plasmid DNA). Amplicon bands of approximately 1.5 kb were expected. A faint band was observed in lane 11, circled in red.

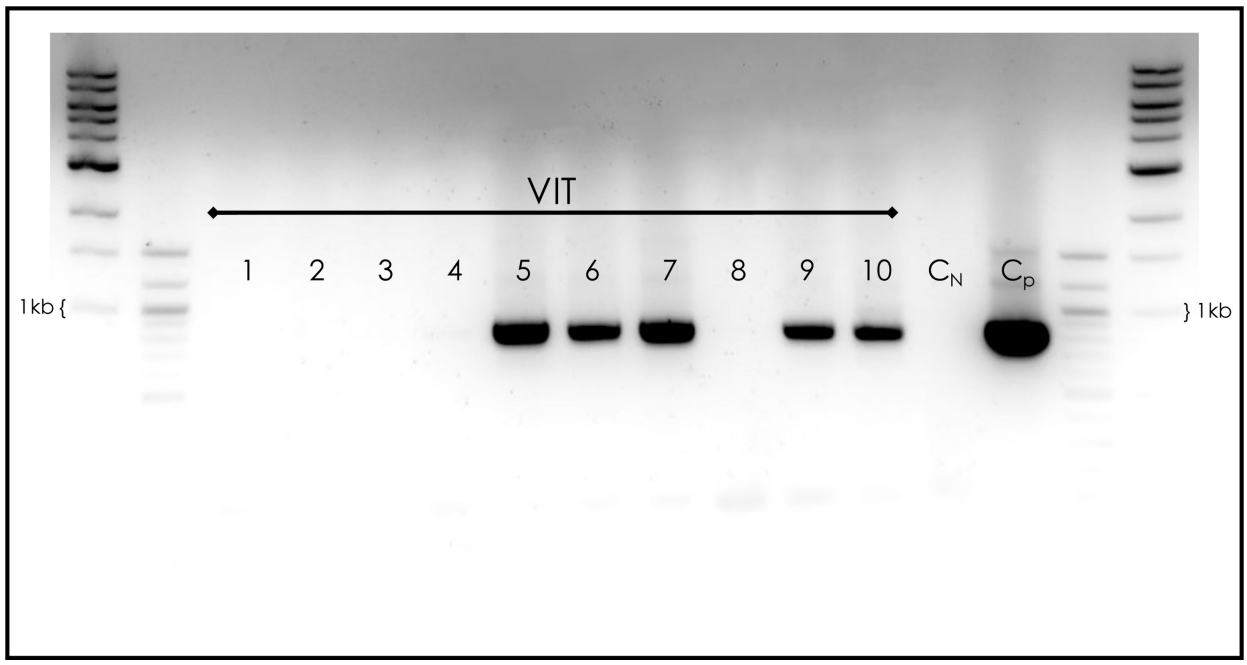


Figure 30. Yeast colony PCR performed on putative positive *S. cerevisiae* (*zrt1zrt2*) transformants thought to contain the pDR195-*VIT* construct. C_N indicates a negative control (i.e. no-template PCR). C_p is a positive control (pDR195-*VIT* plasmid DNA). Amplicon bands of approximately 800 bp were expected.

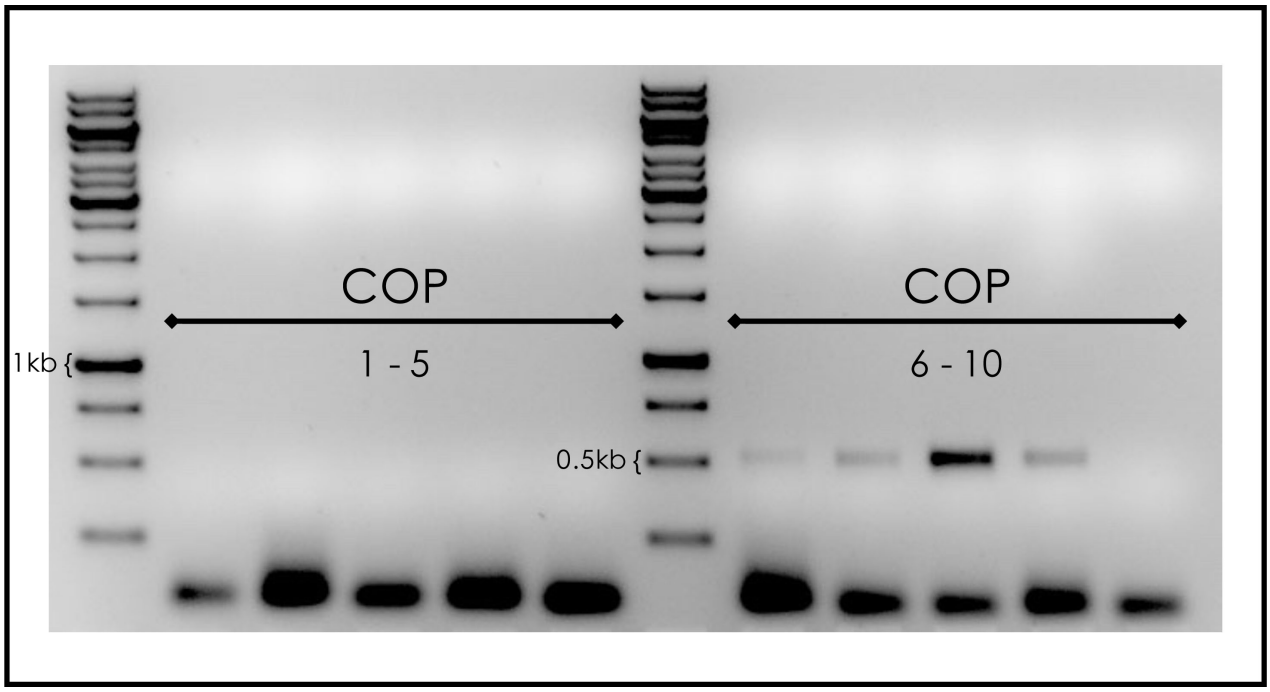


Figure 31. Yeast colony PCR performed on putative positive *S. cerevisiae* (*zrt1zrt2*) transformants thought to contain the pDR195-*COP* construct. Amplicon bands of approximately 500 bp were expected.

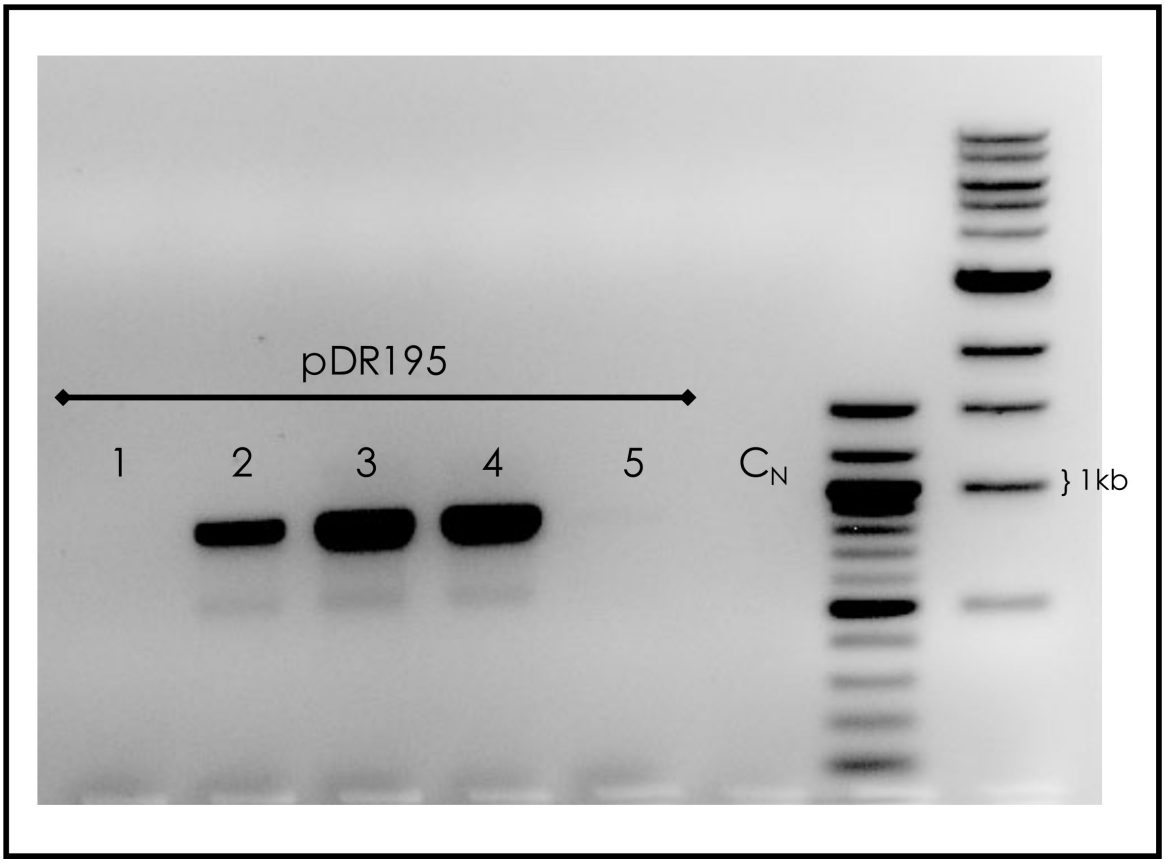


Figure 32. Yeast colony PCR performed on putative positive *S. cerevisiae* (*zrt1zrt2*) transformants thought to contain the empty pDR195 vector. C_N indicates a negative control (i.e. no-template PCR). Amplicon bands of approximately 860 bp were expected.

Nickel sensitivity assays

In order to evaluate the effect of nickel on yeast expressing ScMATE, ScVIT or ScCOP, the impact of nickel on yeast containing the empty pDR195 vector first needed to be established. As such, an initial assay was conducted whereby *zrt1zrt2* transformants containing the empty pDR195 plasmid were cultured on media supplemented with increasing concentrations of nickel. This allowed for the determination of the concentration of nickel at which an inhibitory effect on growth is observed in this strain. CSM-uracil plates were prepared with the following range of nickel chloride concentrations: 0 mM, 1.0 mM, 1.25 mM, 1.5 mM, 1.75 mM and 2.0 mM. Colony growth was not affected up to a concentration of 1.25 mM $NiCl_2$. Colony size began to decrease from 1.5 mM $NiCl_2$ and was almost entirely inhibited at the lower cell dilutions by 2.0 mM $NiCl_2$. See Figure 33.

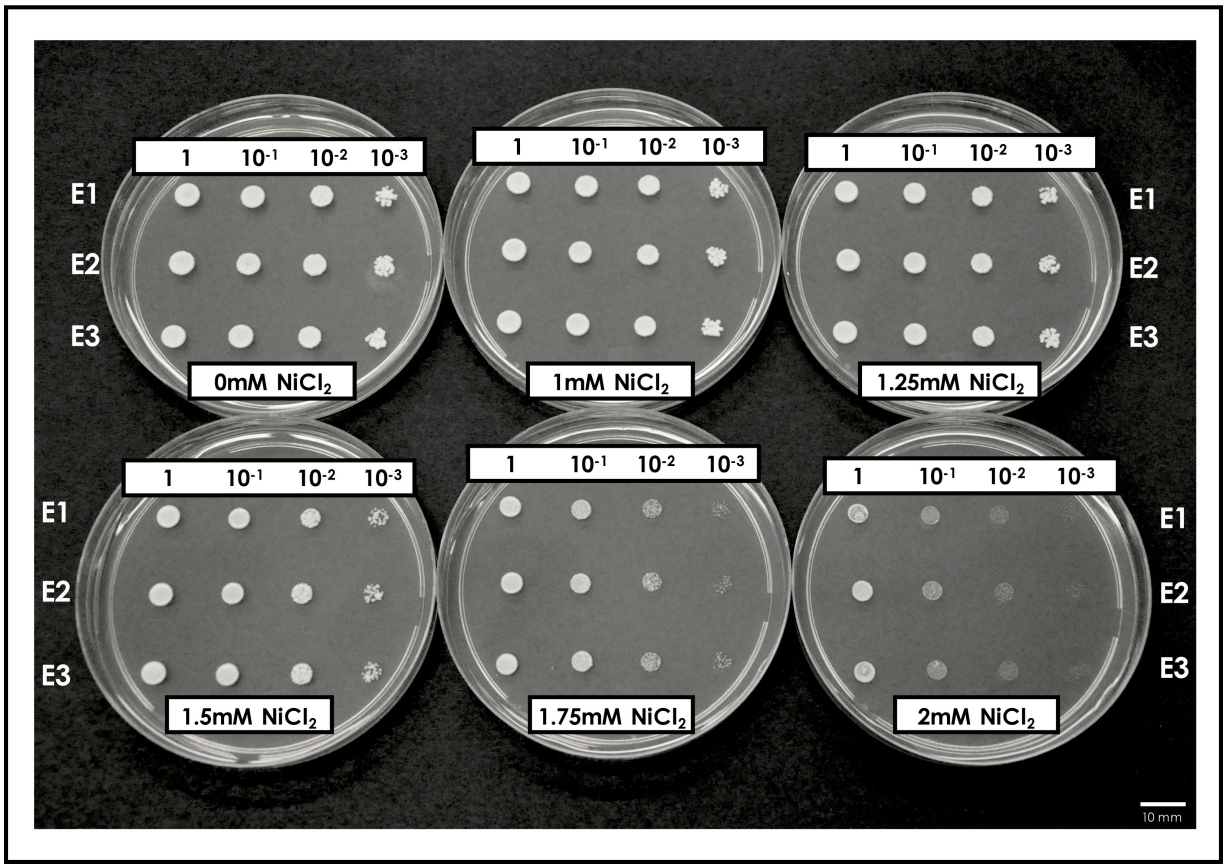


Figure 33. Nickel sensitivity assay in which liquid cultures of confirmed positive *S. cerevisiae* (*zrt1zrt2*) transformants containing the empty pDR195 vector were spotted out at a range of dilutions across six CSM-uracil agar plates containing increasing concentrations of nickel chloride. Three independent transformants were used (designated E1-3). 5 µL of each culture was plated out.

Thus, having established that growth of *zrt1zrt2* cells (containing the empty pDR195 vector) was possible but noticeably inhibited at 1.75 mM $NiCl_2$ and 2 mM $NiCl_2$ these concentrations were selected as the points at which to look for potential effects of ScMATE, ScVIT and ScCOP on nickel activity. As such, the assay was conducted on CSM-uracil plates supplemented with nickel chloride at concentrations of 0 mM, 1.75 mM and 2.0 mM. Yeast transformants containing pDR195-*ScMATE*, pDR195-*ScVIT*, pDR195-*ScCOP*, pDR195-*AtIREG2* as well as the empty pDR195 plasmid were spotted onto these plates at OD_{600} = ±1 and serial dilutions of 10^{-1}, 10^{-2} and 10^{-3}. See Figure 34. The plate not supplemented with nickel chloride was photographed after three days, whereas the nickel-supplemented plates were photographed after four days.

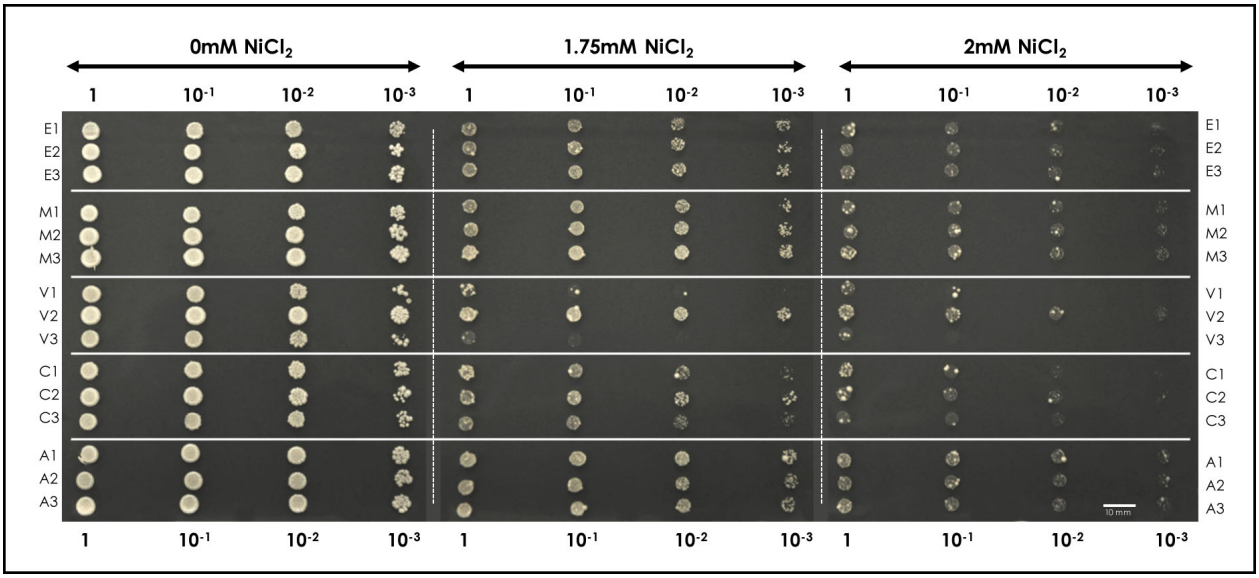


Figure 34. Nickel sensitivity assay in which confirmed positive *S. cerevisiae* (*zrt1zrt2*) transformants containing the empty pDR195 vector (E), pDR195-*ScMATE* (M), pDR195-*ScVIT* (V), pDR195-*ScCOP* (C) and pDR195-*AtIREG2* (A) were spotted out on three CSM-uracil agar plates containing 0 mM, 1.75 mM or 2.0 mM nickel chloride. Three independent transformants were used for each yeast mutant strain and 3 µL of each culture was plated out.

At first glance, there appeared to be very little difference in the colony sizes between the different mutant strains at each nickel concentration. Closer inspection, however, revealed what may be subtle differences between the strains. Nonetheless, the differences observed between the negative (i.e. empty pDR195) and positive (i.e. *AtIREG2*) control colonies are relatively moderate, which limits the strength of

any inferences to be drawn from the experimental data. As such, the following interpretation of these data is considered tentative, if not speculative, until such time as these experiments can be reproduced.[54]

Expression of ScMATE and AtIREG2 appeared to very slightly increase resistance to nickel; colony growth seeming comparatively more robust than that of the negative control. Conversely, colony growth of ScCOP transformants seemed to be inhibited in the media supplemented with $NiCl_2$expression of ScCOP, which could indicate an increase in nickel sensitivity. The apparent increase in nickel sensitivity observed here may imply that ScCOP plays role in nickel transport which, if confirmed, could suggest that ScCOP localises either to the tonoplast (thus functioning to shuttle nickel chelates into the vacuole) or to the plasma membrane (thus functioning as a metal efflux protein). In the case of ScVIT, the growth of clones V1 and V3 was minimal compared to V2 even in the 0 mM $NiCl_2$ media (at the 10^{-3} dilution).[55] However, if one were to speculate that V1 and V3 are representative of typical ScVIT colony growth, it would appear to increase nickel sensitivity. Without more experimental data, however, it would be imprudent to draw conclusions from these limited data. Experimental data regarding nickel accumulation within *zrt1zrt2* yeast expressing ScMATE, ScVIT and ScCOP were not collected here but would be an obvious next step in further elucidating the role of these proteins in nickel transport and accumulation.

The data presented here represent a very tentative first step towards characterising the role of MATE, VIT and COP proteins in *S. coronatus*. Before further experimental work is considered, it would be essential to reproduce this sensitivity assay and confirm the growth pattern observed here. Following that, the obvious next step would be to measure nickel accumulation in all three mutant strains generated here. In a similar study, changes observed in nickel sensitivity in transgenic yeast (expressing

[54] This assay was conducted only once in this study, but was subsequently repeated on two occasions. As such, the experiment was repeated independently on three occasions. The results in the second and third experiments were similar to the first. All three assays can be seen in Appendix 3.
[55] This pattern was observed in the subsequent experiments. It would thus be advisable to repeat the assays with additional transformants.

two NRAMP transporters) were not correlated with any changes in intracellular nickel content (Merlot et al. 2014). Establishing whether heterologous expression of *ScMATE*, *ScVIT* and *ScCOP* affects nickel accumulation in yeast is therefore crucial in the characterising of protein function.

Related to this, although perhaps more critical, was the finding that PgIREG1 localises to the tonoplast when expressed in *A. thaliana* and increased nickel resistance therein (Merlot et al. 2014). This suggests that PgIREG1 acts to transport nickel across the tonoplast into the vacuole, rather than acting as a true exporter by the extrusion of nickel from the plant cell itself. This highlights the risk in making direct, unqualified comparisons between organisms with such different biochemical and metabolic infrastructure as plants and fungi. This finding therefore suggests another important next step in the functional characterisation of these proteins: the *in planta* expression of *ScMATE*, *ScVIT* and *ScCOP* in *A. thaliana* and determination of the subcellular localisation thereof.

Lastly, although post-translational modification (particularly with respect to protein-folding) is considered to be broadly analogous between fungi and plants (Han & Yu 2015), the utility of heterologous expression of plant proteins in *S. cerevisiae* is acknowledged to be limited by both the rate of expression (Han & Yu 2015; Yesilirmak & Sayers 2009) and the "mislocalisation" of membrane proteins (Yesilirmak & Sayers 2009). Thus, although it is possible that these proteins may simply play no role in nickel transport, other factors could very well be at play. It is recommended that these experiments be repeated, if only to establish the consistency of these results, but the limitations of such experiments should not be ignored.

4. Subcellular protein localisation

In Section 2, it was suggested that protein localisation predictions based purely on amino acid sequences are unreliable without supporting experimental data. Such predictions based on the ScMATE and ScVIT sequences indicate targeting to the plasma membrane by ScMATE and either the tonoplast or plasma membrane by ScVIT. Experimental evidence suggests that, in other plant taxa, MATE proteins localise to the plasma membrane (Magalhaes et al. 2007; Perea-García et al. 2013) and VIT proteins localise to the tonoplast (Kim et al. 2006; Zhang et al. 2012; Zhu et al. 2016). Here, fusion proteins comprised of ScMATE or ScVIT with C-terminal EYFP tags were expressed in live onion cells in order to determine their subcellular localisation, thus allowing validation of the sequence-based prediction platforms. Unfortunately, the time available was insufficient for this to be repeated with ScCOP.

PCR amplification, expression vector preparation and cloning

Full-length *MATE* and *VIT* ORFs (minus stop codon) were successfully amplified from *S. coronatus* cDNA (Figure 35). Both *MATE* and *VIT* were ligated into the entry vector pENR1A, then transformed into *E. coli*. Positive transformants were confirmed through colony PCR with gene-specific primers (Figure 36 and 37). The expression vector pG101 was digested with *EcoRI* to confirm its identity. The band sizes observed after electrophoresis were as expected (Figure 38) and confirmed plasmid identity. The *MATE* and *VIT* ORFs were Gateway cloned into pG101 (downstream of the *CaMV 35S* promotor and upstream of the *EYFP* coding sequence), generating the recombinant plasmids *pG101-MATE-EYFP* and *pG101-VIT-EYFP*. The junction between *MATE* or *VIT* and *EYFP* was sequenced and aligned against the validated MATE and VIT sequences and the EYFP sequence (Figures 39 and 40). The alignments confirmed that the reading frame had been conserved, thus translation was expected to result in the expression of in-frame EYFP-fusion proteins. These recombinant plasmids were transformed into *E. coli* and colony PCR (with gene-specific primers) was used to confirm the presence of the GOIs (Figure 41). After propagation in *E. coli*, the extracted plasmids were transformed into *A. tumefaciens* (GV3101), where the presence of the GOI was again confirmed by colony PCR (Figures 42 and 43).

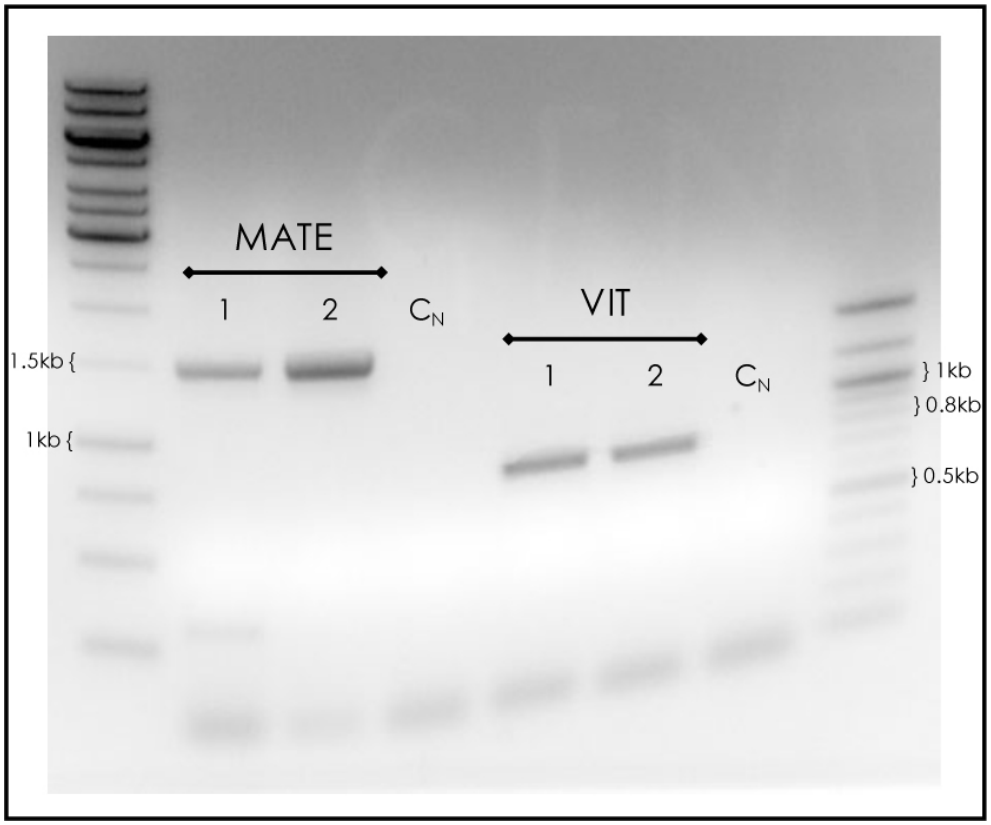


Figure 35. PCR amplification of *MATE* and *VIT* electrophoresed on a 1% (w/v) agarose gel. C_N indicates a negative control (i.e. no-template PCR). Amplicon bands of approximately 1.5 kb and 800 bp respectively were expected.

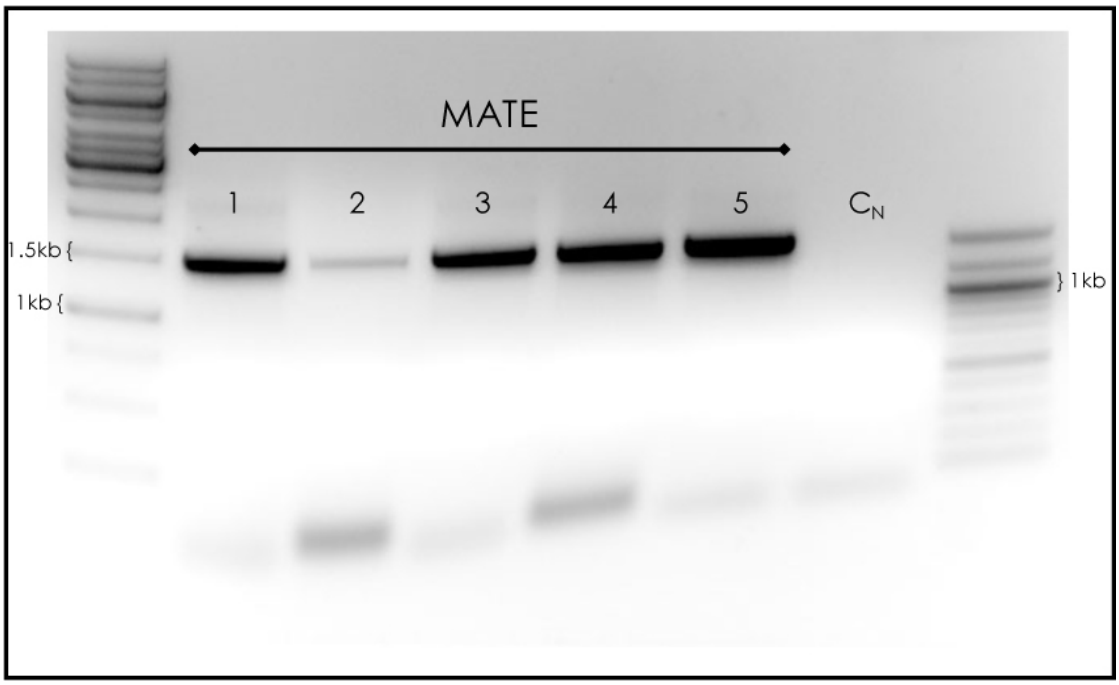


Figure 36. Colony PCR performed on putative positive *E. coli* (DH5α) transformants thought to contain the *pENTR1A-MATE* construct. C_N indicates a negative control (i.e. no-template PCR). Amplicon bands of approximately 1.5 kb were expected.

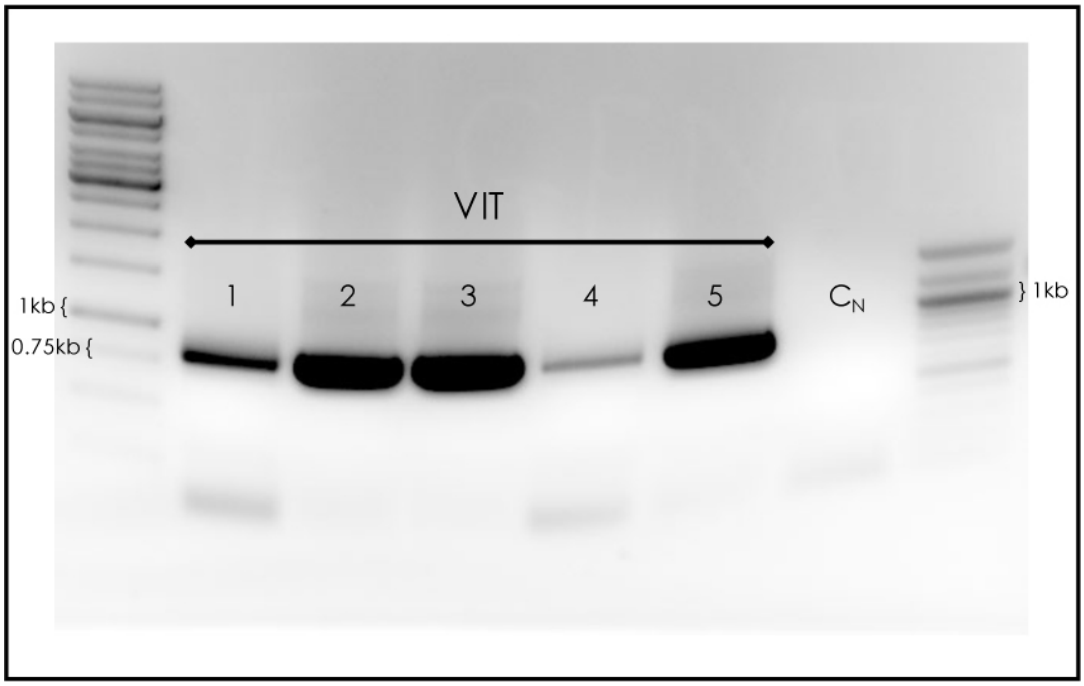


Figure 37. Colony PCR performed on putative positive *E. coli* (DH5α) transformants thought to contain the *pENTR1A-VIT* construct. C_N indicates a negative control (i.e. no-template PCR). Amplicon bands of approximately 800 bp were expected.

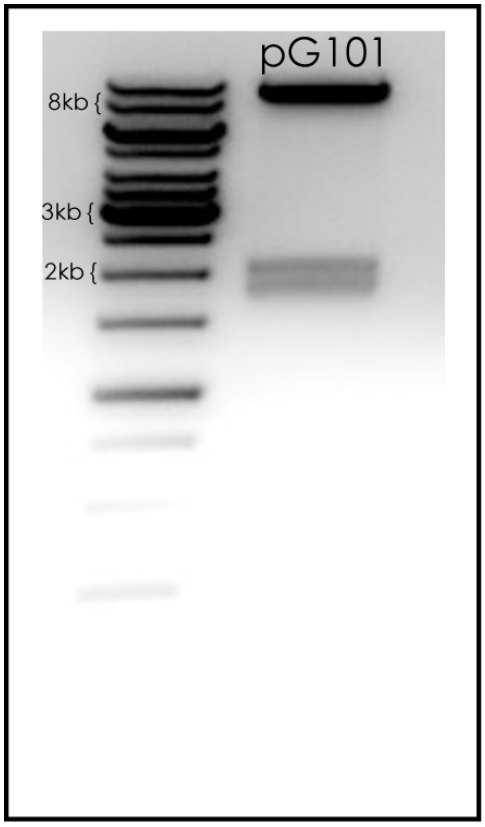


Figure 38. Restriction enzyme digestion of pG101 using *EcoRI*. Fragments of 8652 bp, 2009 bp and 1792 bp were expected.

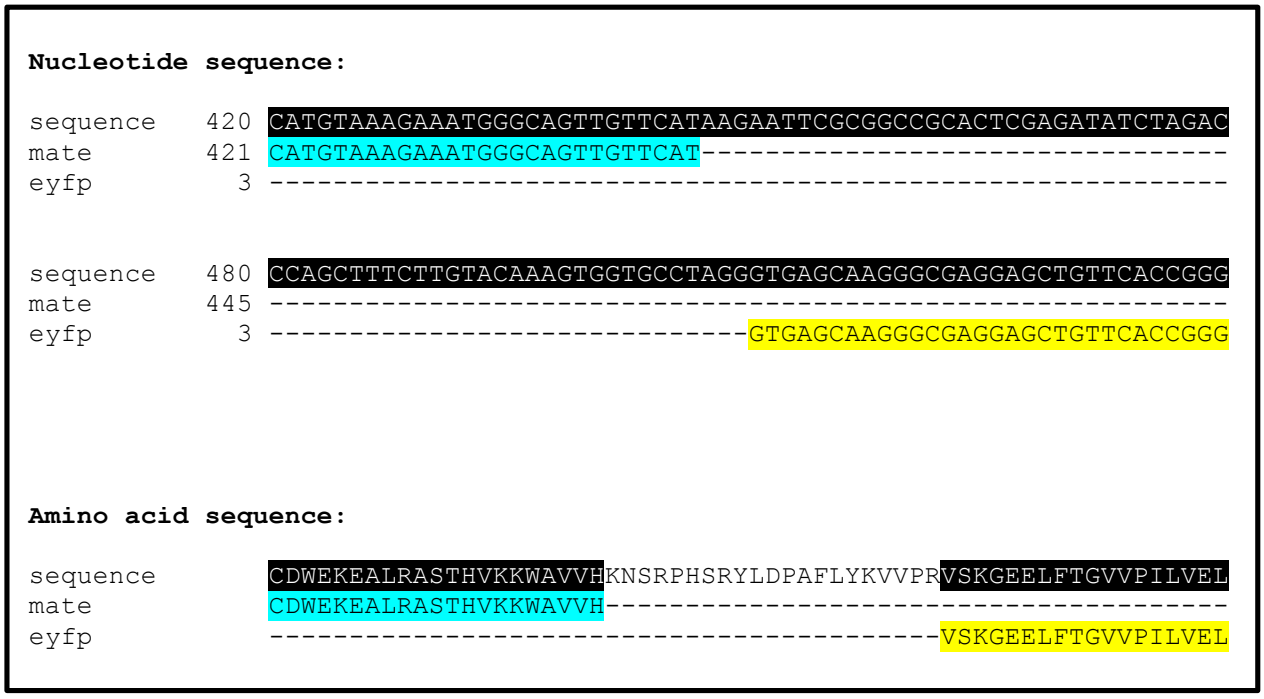


Figure 39. Partial nucleotide (top) and amino acid (bottom) sequence alignments of the *pG101-MATE-EYFP* sequence ("sequence"), the validated MATE sequence ("mate", highlighted in blue) and the EYFP sequence ("eyfp", highlighted in yellow).

Figure 40. Partial nucleotide (top) and amino acid (bottom) sequence alignments of the *pG101-VIT-EYFP* sequence ("sequence"), the validated VIT sequence ("vit", highlighted in pink) and the EYFP sequence ("eyfp", highlighted in yellow).

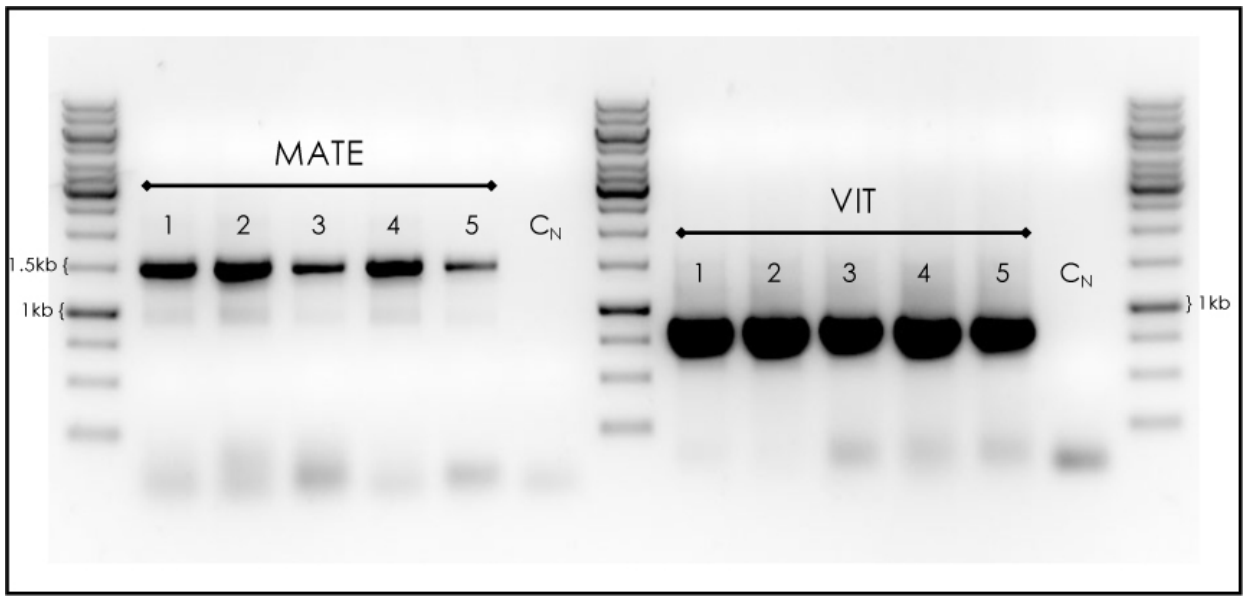


Figure 41. Colony PCR performed on putative positive *E. coli* (DH5α) transformants thought to contain the *pG101-MATE-EYFP* and *pG101-VIT-EYFP* constructs. C_N indicates a negative control (i.e. no-template PCR). Amplicon bands of approximately 1.5 kb and 800 bp respectively were expected.

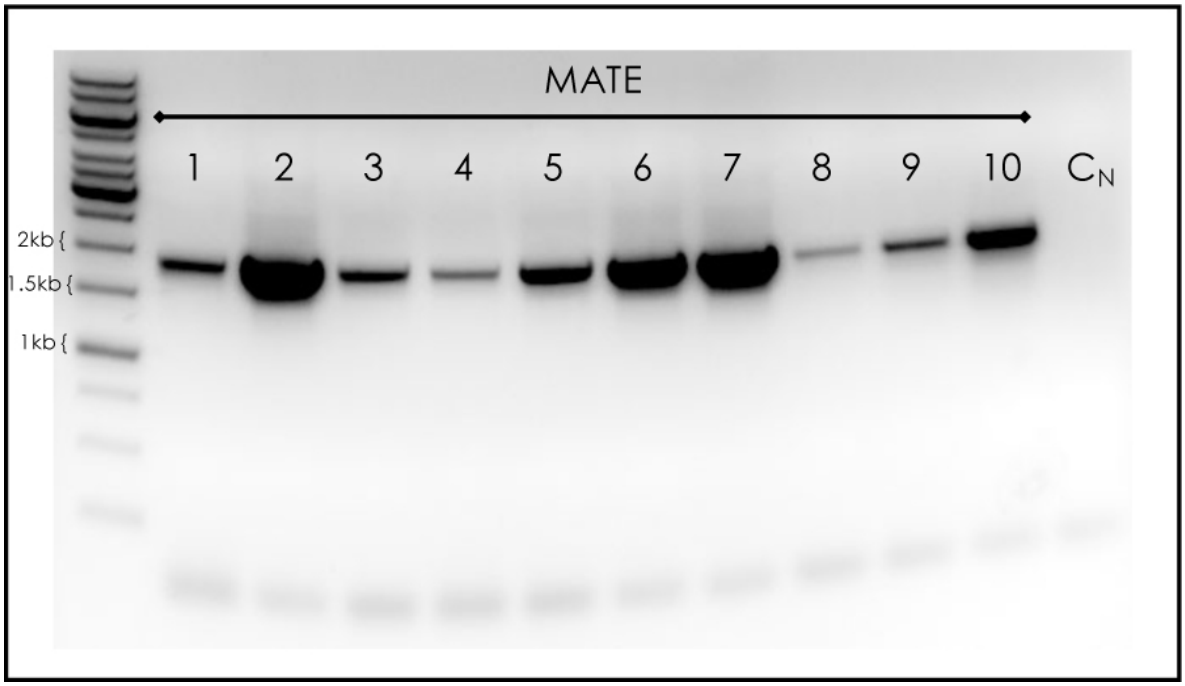


Figure 42. Colony PCR performed on putative positive *A. tumefaciens* (GV3101) transformants thought to contain the *pG101-MATE-EYFP* construct. C_N indicates a negative control (i.e. no-template PCR). Amplicon bands of approximately 1.5 kb were expected.

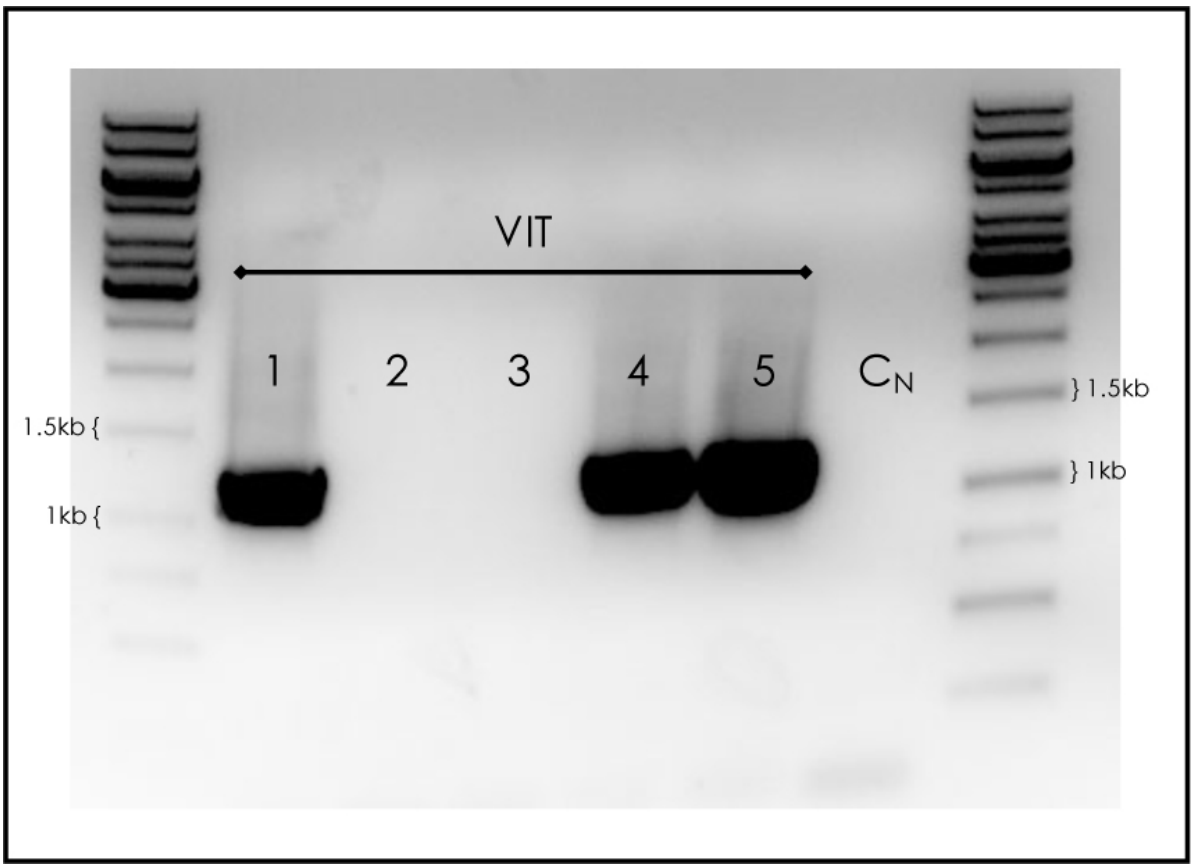


Figure 43. Colony PCR performed on putative positive *A. tumefaciens* (GV3101) transformants thought to contain the *pG101-VIT-EYFP* construct. C_N indicates a negative control (i.e. no-template PCR). Amplicon bands of approximately 800 bp were expected.

Fusion protein localisation in onion epidermal cells

The expression of fluorescence-tagged proteins in living cells by transient transformation methods commonly involves particle bombardment (Ueki et al. 2009; Zhang et al. 2003), protoplast transfection (Chen et al. 2006) or various *Agrobacterium*-mediated transformation techniques (Li et al. 2009; Marion et al. 2008; Ye et al. 1999). A modified *Agrobacterium*-mediated transient transformation method has been shown to improve transformation efficiency by up to an order of magnitude when compared to other methods (Xu et al. 2014) and was used here to transfect live onion cells with the recombinant plasmids *pG101-MATE-EYFP* and *pG101-VIT-EYFP*. Following *Agrobacterium* infiltration, the subcellular localisation of MATE and VIT fusion proteins was analysed using fluorescence microscopy.

Negative controls consisted of both untransformed onion epidermis (Figure 44A and 44B) and control transfections conducted with untransformed *A. tumefaciens* (Figure 44C and 44D). The autofluorescence signals were minimal in both controls, but serve to illustrate the background levels of fluorescence expected. *A. tumefaciens* transformed with the empty pG101 vector was not used in control experiments; the C-terminal *EYFP* tag lacks a start codon thus would not be expressed in unfused form.

In Figures 45 and 46, the overlay of bright-field and fluorescence images strongly suggests that MATE-EYFP localises to the nucleus. Figures 46-48 are less readily susceptible to interpretation. In Figures 47 and 48 (VIT-EYFP), the fluorescent signal is less easy to assign to a specific area or organelle within the plant cell, although the localisation to the tonoplast, plasma membrane or even the cytoplasm cannot be excluded. Aspects of both figures resemble the confocal micrographs (Figure 49) of an IREG1 protein described as localising to the tonoplast in a study of another nickel hyperaccumulator (Merlot et al. 2014). Without further experimental data, however, it would be difficult to make any firm deductions regarding VIT-EYFP localisation. Expression of fusion proteins known to localise to the tonoplast and plasma membrane would allow a more direct comparison against the signal observed here.

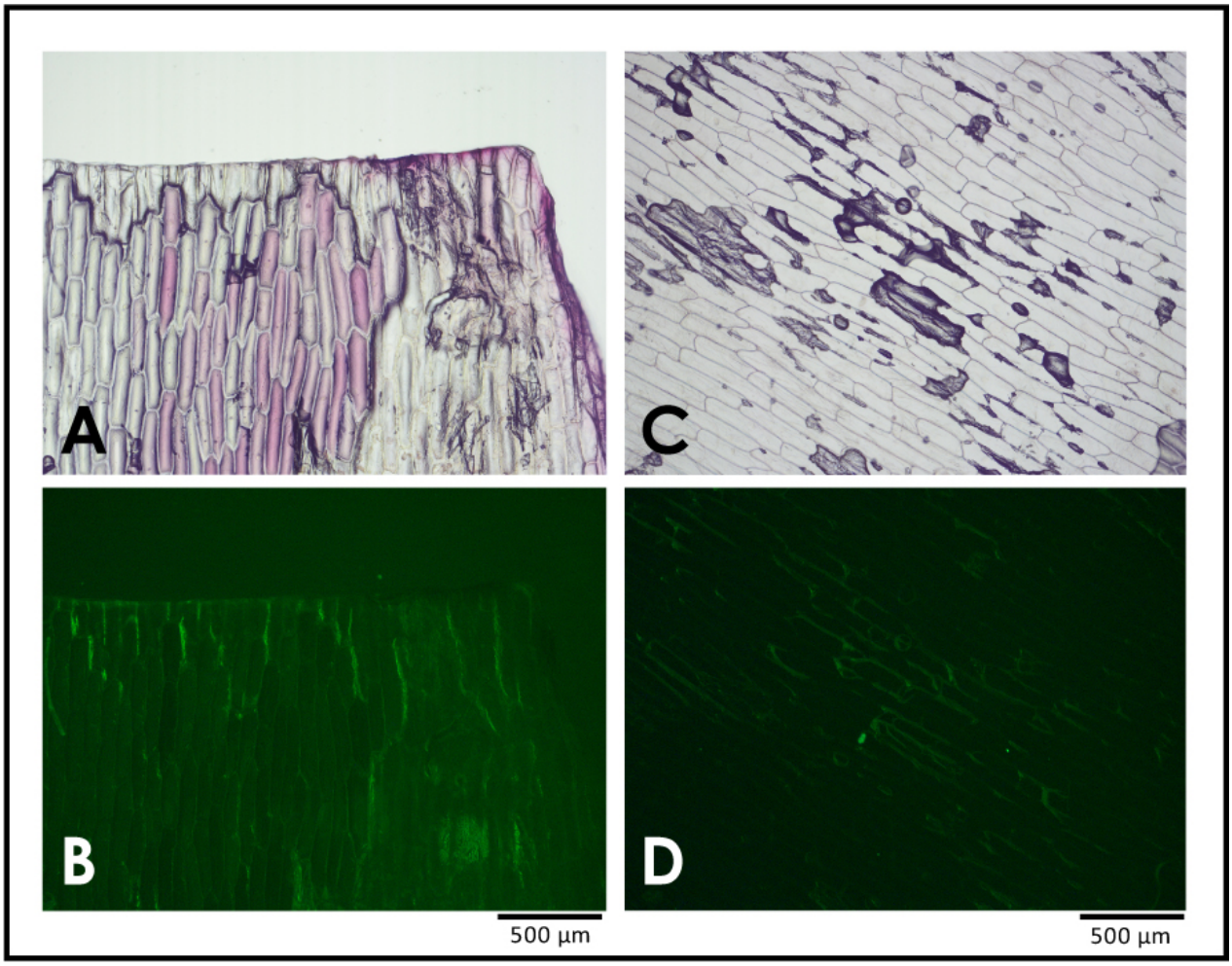


Figure 44. Micrographs (4x magnification) of onion cells of untransfected onion cells (**A**. Bright-field image. **B**. Fluorescence image.) and onion cells infiltrated with untransformed *A. tumefaciens* (**C**. Bright-field image. **D**. Fluorescence image.).

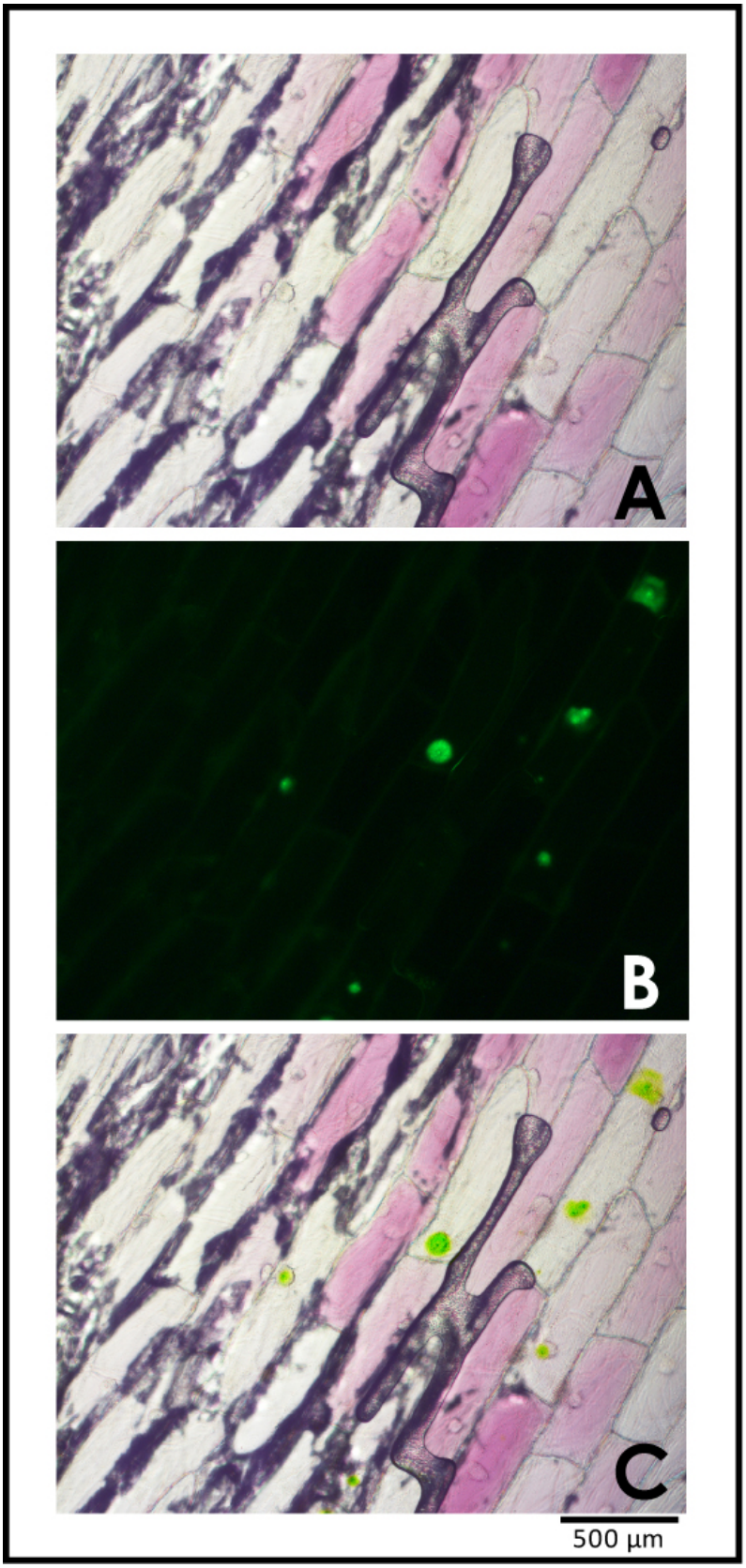


Figure 45. *Agrobacterium*-mediated transient transfection of onion cells. Micrographs (4x magnification) of onion cells transfected with *pG101-MATE-EYFP*. **A**. Bright-field image. **B**. Fluorescence image. **C**. Image overlay.

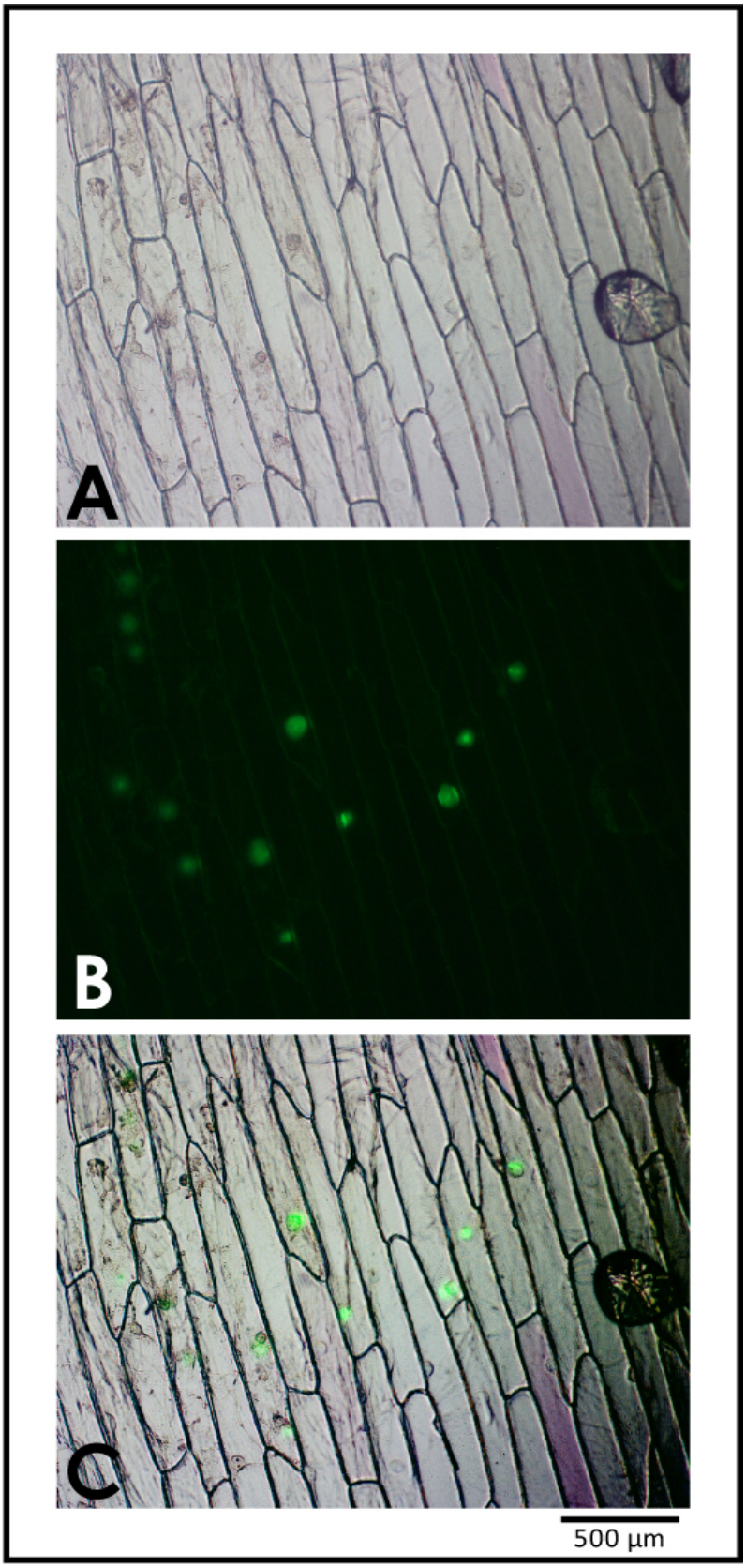


Figure 46. *Agrobacterium*-mediated transient transfection of onion cells. Micrographs (4x magnification) of onion cells transfected with *pG101-MATE-EYFP*. **A**. Bright-field image. **B**. Fluorescence image. **C**. Image overlay.

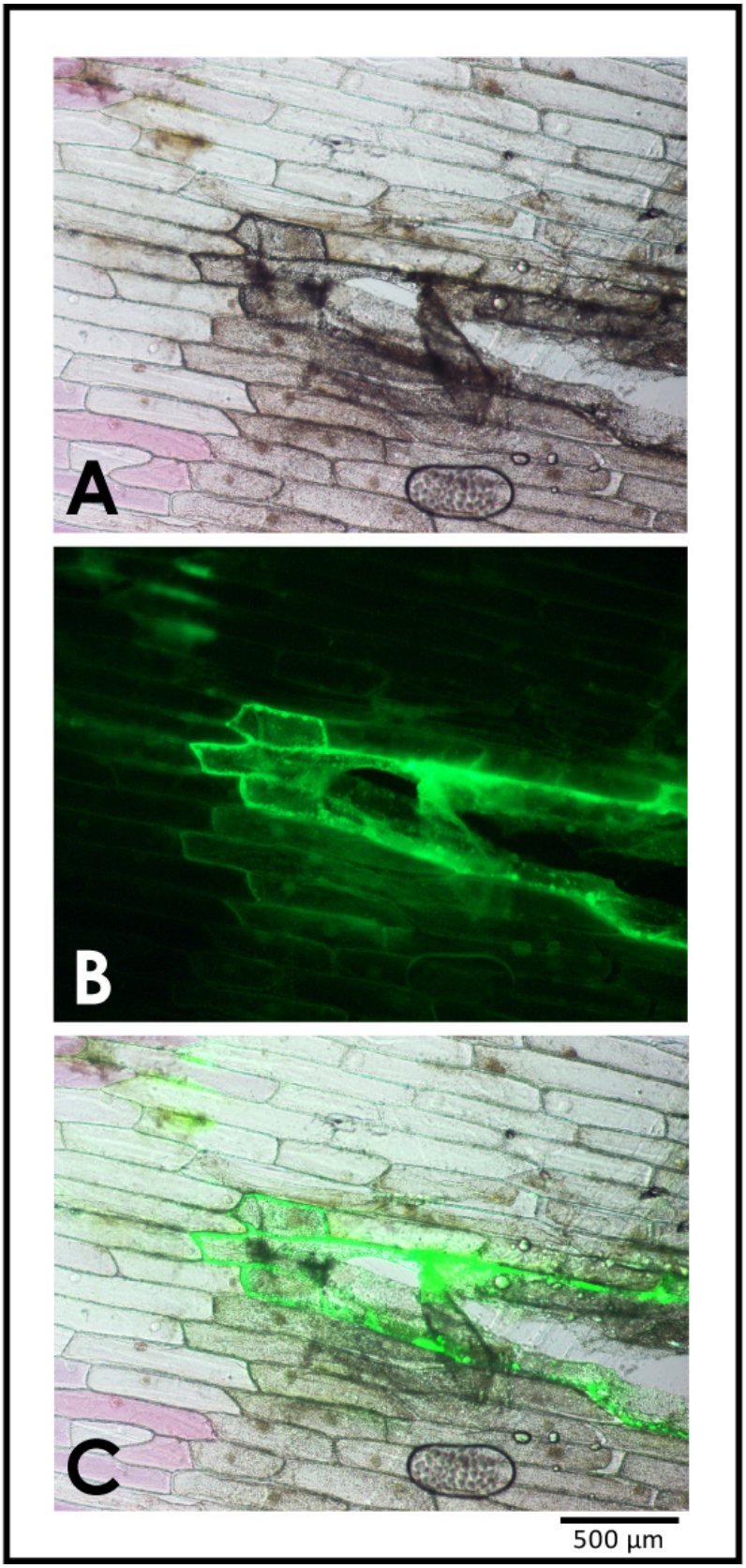


Figure 47. *Agrobacterium*-mediated transient transfection of onion cells. Micrographs (4x magnification) of onion cells transfected with *pG101-VIT-EYFP*. **A**. Bright-field image. **B**. Fluorescence image. **C**. Image overlay.

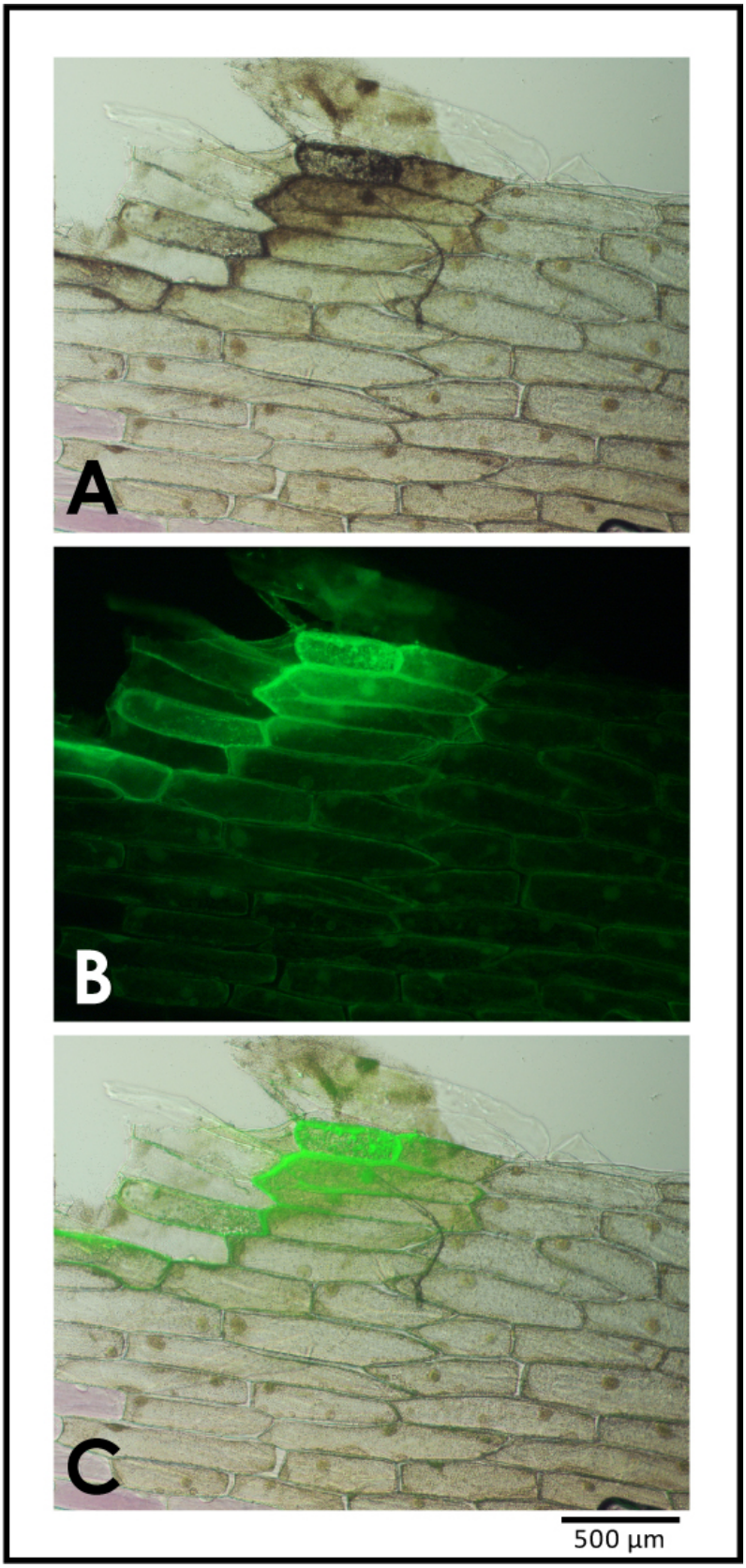


Figure 48. *Agrobacterium*-mediated transient transfection of onion cells. Micrographs (4x magnification) of onion cells transfected with *pG101-VIT-EYFP*. **A**. Bright-field image. **B**. Fluorescence image. **C**. Image overlay.

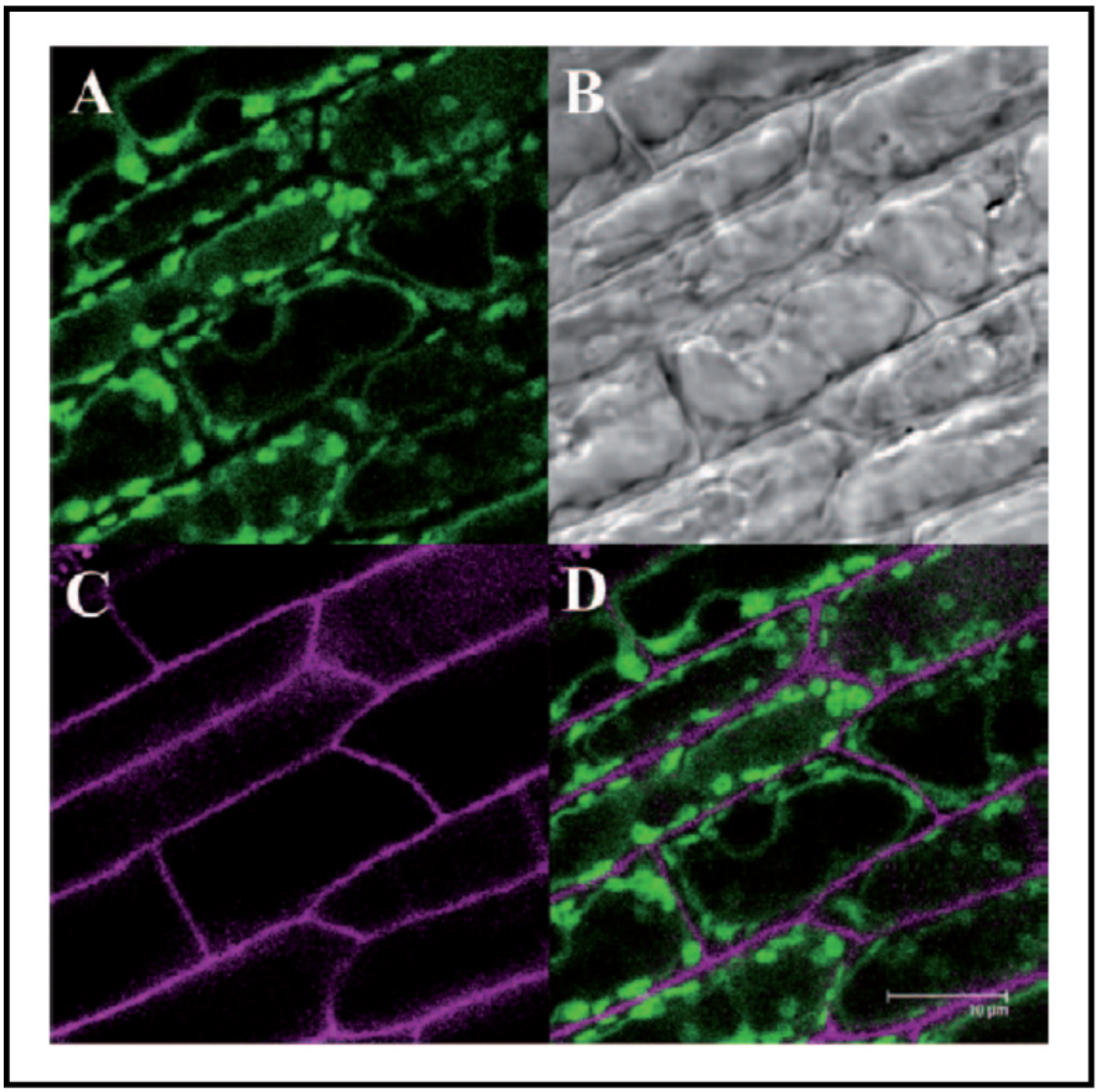


Figure 49. Confocal micrographs (scale bar = 10 µm) of plant cells of expressing a GFP-tagged tonoplast nickel transport protein, PgIREG1. **A**. Fluorescence image (GFP). **B**. Bright-field image. **C**. Fluorescence image (propidium iodide used to stain cell walls). **D**. Image overlay. Image from Merlot et al. 2014.

To date, MATE family proteins have not been demonstrated to localise to the nucleus. A first step in the further investigation of this unusual result would be to repeat the experiment using a fluorescent nuclear stain such as Hoechst or DAPI to more clearly delineate the nucleus during microscopy. The apparent localisation of VIT-EYFP is less surprising, although nonetheless requiring of further investigation. Both experimental evidence and amino acid sequence-based predictions suggest that VIT proteins target the tonoplast. Although this possibility is by no means eliminated here, further experimental work involving vacuolar staining or the imaging of cells expressing confirmed tonoplast proteins would provide clarity in this regard. Such an approach would also conclusively exclude any possibility that the fluorescent signal detected here is caused by autofluorescence.

Thus, although no firm conclusions can currently be drawn from these data, further experimental work would serve to support and/or refine these findings. A first step in this regard would be to repeat the experiment using N-terminal EYFP fusion proteins, as differences in protein localisation have been observed between otherwise identical proteins tagged with either N- or C-terminal fluorescent markers (Merlot et al. 2014). This would further allow for the *in planta* expression of unfused EYFP, which would indicate typical cytoplasmic EYFP localisation.

IV GENERAL SYNTHESIS

The ability of hyperaccumulator plants to amass astonishingly high concentrations of intracellular heavy/transition metals is an extreme trait eminently deserving of investigation. Here, genes identified as being upregulated in hyperaccumulating populations of *S. coronatus* (compared to their non-accumulating conspecifics) were investigated with respect to a possible role in nickel transport.

The first objective was to validate the sequences derived from prior RNA-Seq work, in which *ScMATE*, *ScVIT* and *ScCOP* were identified as possibly encoding transport proteins. Nucleotide sequences generated from cDNA were translated and these amino acid sequences aligned perfectly with those translated from the RNA-Seq sequences for *ScVIT* and *ScCOP*. The amino acid alignment for ScMATE, however, revealed a number of substitutions. It is recommended that further individuals from different populations be sequenced in order to determine the validity of the observed changes.

Bioinformatics-based predictions of transmembrane helices in the proteins encoded by these genes were remarkably similar to those for their *A. thaliana* homologs. In *A. thaliana*, MATE, VIT and COP proteins are predicted to have twelve, five and three transmembrane domains respectively (Braibant et al. 2002; Brown et al. 1999; Maron et al. 2010; Puig & Thiele 2002; Shitan et al. 2014; Slavic et al. 2016; Zhang et al. 2012). These predictions were replicated for ScMATE, ScVIT and ScCOP. Protein localisation predictions, although considered unreliable, suggested that ScMATE might localise to the plasma membrane, ScVIT to the tonoplast and ScCOP to the plasma membrane or tonoplast.

In order to determine whether ScMATE, ScVIT and ScCOP are capable of transporting nickel, transgenic, metal-sensitive yeast strains were cultured on media containing increasing concentrations of nickel. Although differences in

colony growth were not substantial, it did appear as if *AtIREG2*-transformed cultures were slightly less sensitive to nickel than the negative control (empty pDR195), which has been observed in other studies (Merlot et al. 2014). ScMATE colonies possibly did slightly better than the negative control, ScCOP colonies slightly worse. This would suggest that the former may possibly be functioning as nickel exporters, the latter as a nickel importer. Unfortunately, time constraints prevented the replication of this assay and, as such, undue reliance should not be placed on these results. An essential next step would be the replication of this assay, perhaps at a greater range of nickel concentrations. Another experiment, imperative to characterising the roles of these proteins, would be to measure the accumulation of nickel in these yeast stains grown in nickel-supplemented liquid media. Thus, the data presented here are considered preliminary.

The EYFP-tagged ScMATE was, surprisingly, shown to localise to the nucleus when expressed in onion cells. This was not consistent with any of the localisation predictions generated across several platforms, nor has such a localisation been observed for MATE family proteins in other taxa. The targeting of ScVIT cannot be adequately determined on the basis of the data presented here, but neither the tonoplast nor the plasma membrane can be excluded. It is recommended that similar work be performed using N-terminal fusion proteins, both C- and N-terminal fusion proteins with known localisations, and that organelle-specific dyes be used to delineate the boundaries of cell structures. Other future work worth pursuing would be the creation of transgenic *A. thaliana* lines expressing fluorescence-tagged *S. coronatus* proteins as well a testing nickel tolerance of plants expressing these genes.

Although the results presented here are not conclusive, they do support a finding that *ScMATE*, *ScVIT* and *ScCOP* should be investigated further as candidate genes for nickel tolerance in *S. coronatus*.

V REFERENCES

Anderson, C. W. N., Brooks, R. R., Chiarucci, A., Lacoste, C. J., Leblanc, M., Robinson, B. H., Simcock, R., & Stewart, R. B. (1999). 'Phytomining for nickel, thallium and gold.' *Journal of Geochemical Exploration*, 67/1–3: 407–15.

Antonovics, J., Bradshaw, A. D., & Turner, R. G. (1971). 'Heavy Metal Tolerance in Plants.' *Advances in Ecological Research*, 7: 1–85.

Assunção, A. G. L., Da Costa Martins, P., De Folter, S., Vooijs, R., Schat, H., & Aarts, M. G. M. (2001). 'Elevated expression of metal transporter genes in three accessions of the metal hyperaccumulator *Thlaspi caerulescens*.' *Plant, Cell and Environment*, 24/2: 217–26.

Baker, A. J. M., & Brooks, R. R. (1989). 'Terrestrial higher plants which hyperaccumulate metallic elements - a review of their distribution, ecology and phytochemistry.' *Biorecovery*, 1/2: 81–126.

Becher, M., Talke, I. N., Krall, L., & Krämer, U. (2004). 'Cross-species microarray transcript profiling reveals high constitutive expression of metal homeostasis genes in shoots of the zinc hyperaccumulator *Arabidopsis halleri*.' *The Plant Journal*, 37/2: 251–68.

Bhargava, A., Carmona, F. F., Bhargava, M., & Srivastava, S. (2012). 'Approaches for enhanced phytoextraction of heavy metals.' *Journal of Environmental Management*, 105: 103–20.

Boyd, R. S. (2007). 'The defense hypothesis of elemental hyperaccumulation: status, challenges and new directions.' *Plant and Soil*, 293/1–2: 153–76.

Boyd, R. S., Davis, M. A., Wall, M. A., & Balkwill, K. (2002). 'Nickel defends the South African hyperaccumulator *Senecio coronatus* (Asteraceae) against *Helix aspersa* (Mollusca: Pulmonidae).' *American Journal of Botany*, 97: 91–7.

Boyd, R. S., & Martens, S. N. (1992). 'The raison d'être for metal hyperaccumulation by plants.' Baker A. J. M., Proctor J., & Reeves R. D. (eds) *The Vegetation of Ultramafic (Serpentine) Soils: Proceedings of the First International Conference on Serpentine Ecology*, pp. 279–89. Andover: Hampshire.

Braibant, M., Guilloteau, L., & Zygmunt, M. S. (2002). 'Functional characterization of *Brucella melitensis* NorMI, an efflux pump belonging to the multidrug and toxic

compound extrusion family.' *Antimicrobial Agents and Chemotherapy*, 46/9: 3050–3.

Brooks, R., Anderson, C., Stewart, R., & Robinson, B. (1999). 'Phytomining: growing a crop of a metal.' *Biologist*, 46/5: 201–5.

Brooks, R., Lee, J., Reeves, R., & Jaffre, T. (1977). 'Detection of nickeliferous rocks by analysis of herbarium specimens of indicator plants.' *Journal of Geochemical Exploration*, 7: 49–57.

Brown, M. H., Paulsen, I. T., & Skurray, R. A. (1999). 'The multidrug efflux protein NorM is a prototype of a new family of transporters.' *Molecular Microbiology*, 31/1: 394–5.

Caesalpino, A. (1583). *De Plantis Libri*. Florentiae: Florentiae.

Callahan, D. L., Baker, A. J. M., Kolev, S. D., & Wedd, A. G. (2006). 'Metal ion ligands in hyperaccumulating plants.' *Journal of Biological Inorganic Chemistry*, 11/1: 2–12.

Chaney, R. L., Malik, M., Li, Y. M., Brown, S. L., Brewer, E. P., Angle, J. S., & Baker, A. J. M. (1997). 'Phytoremediation of soil metals.' *Current Opinion in Biotechnology*, 8/3: 279-84.

Chang, T. H., Wu, L. C., Lee, T. Y., Chen, S. P., Huang, H. D., & Horng, J. T. (2013). 'EuLoc: a web-server for accurately predict protein subcellular localization in eukaryotes by incorporating various features of sequence segments into the general form of Chou's PseAAC.' *Journal of Computer-Aided Molecular Design*, 27/1: 91–103.

Chen, S., Tao, L., Zeng, L., Vega-Sanchez, M. E., Umemura, K., & Wang, G. L. (2006). 'A highly efficient transient protoplast system for analyzing defence gene expression and protein-protein interactions in rice.' *Molecular Plant Pathology*, 7/5: 417–27.

Chomczynski, P., & Sacchi, N. (1987). 'Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction.' *Analytical Biochemistry*, 162/1: 156–9.

Christenhusz, M. J. M., & Byng, J. W. (2016). 'The number of known plants species in the world and its annual increase.' *Phytotaxa*, 261/3: 201–17.

Clemens, S. (2006). 'Toxic metal accumulation, responses to exposure and mechanisms of tolerance in plants.' *Biochimie*, 88/11: 1707–19.

Clemens, S., Palmgren, M. G., & Krämer, U. (2002). 'A long way ahead: understanding and engineering plant metal accumulation.' *Trends in Plant Science*, 7/7: 309–15.

Cvitanich, C., Przybyłowicz, W. J., Urbanski, D. F., Jurkiewicz, A. M., Mesjasz-Przybyłowicz, J., Blair, M. W., Astudillo, C., Jensen, E., & Stougaard, J. (2010). 'Iron and ferritin accumulate in separate cellular locations in *Phaseolus* seeds.' *BMC Plant Biology*, 10/1: 26.

DalCorso, G., Manara, A., Piasentin, S., & Furini, A. (2014). 'Nutrient metal elements in plants.' *Metallomics*, 6/10: 1770–88.

Diener, A. C., Gaxiola, R. A., & Fink, G. R. (2001). '*Arabidopsis* ALF5, a multidrug efflux transporter gene family member, confers resistance to toxins.' *Plant Cell*, 13/7: 1625–38.

Felsenstein, J. (1985). 'Confidence limits on phylogenies: an approach using the bootstrap.' *Evolution*, 39/4: 783.

Fones, H., Davis, C. A. R., Rico, A., Fang, F., Smith, J. A. C., & Preston, G. M. (2010). 'Metal hyperaccumulation armors plants against disease.' *PLoS Pathogens*, 6/9: 1-13.

Gartside, D. W., & McNeilly, T. (1974). 'Genetic studies in heavy metal tolerant plants.' *Heredity*, 33/3: 303–8.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. (2003). 'ExPASy: the proteomics server for in-depth protein knowledge and analysis.' *Nucleic Acids Research*, 31/13: 3784–8.

Gendre, D., Czernic, P., Conéjéro, G., Pianelli, K., Briat, J. F., Lebrun, M., & Mari, S. (2007). '*TcYSL3*, a member of the *YSL* gene family from the hyper-accumulator *Thlaspi caerulescens*, encodes a nicotianamine-Ni/Fe transporter.' *Plant Journal*, 49/1: 1–15.

Gollhofer, J., Timofeev, R., Lan, P., Schmidt, W., & Buckhout, T. J. (2014). 'Vacuolar-iron-transporter1-like proteins mediate iron homeostasis in *Arabidopsis*.' *PLoS ONE*, 9/10: 1–8.

Goolsby, E. W., & Mason, C. M. (2015). 'Toward a more physiologically and evolutionarily relevant definition of metal hyperaccumulation in plants.' *Frontiers in Plant Science*, 6/33: 1-4.

Han, M., & Yu, X. (2015). 'Enhanced expression of heterologous proteins in yeast cells via the modification of N-glycosylation sites.' *Bioengineered*, 6/2: 115–8.

Haydon, M. J., & Cobbett, C. S. (2007). 'Transporters of ligands for essential metal ions in plants.' *New Phytologist*, 174/3: 499–506.

Holsters, M., Silva, B., Van Vliet, F., Genetello, C., De Block, M., Dhaese, P., Depicker, A., Inzé, D., Engler, G., Villarroel, R., Van Montagu, M., & Schell, J. (1980). 'The functional organization of the nopaline *A. tumefaciens* plasmid pTiC58.' *Plasmid*, 3/2: 212–30.

Hooper, C. M., Tanz, S. K., Castleden, I. R., Vacher, M. A., Small, I. D., & Millar, A. H. (2014). 'SUBAcon: a consensus algorithm for unifying the subcellular localization data of the *Arabidopsis* proteome.' *Bioinformatics*, 30/23: 3356–64.

Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., & Nakai, K. (2007). 'WoLF PSORT: protein localization predictor.' *Nucleic Acids Research*, 35: 585-7.

Huitson, S. B., & Macnair, M. R. (2003). 'Does zinc protect the zinc hyperaccumulator *Arabidopsis halleri* from herbivory by snails?' *New Phytologist*, 159/2: 453–9.

Ingle, R. A., Mugford, S. T., Rees, J. D., Campbell, M. M., & Smith, J. A. C. (2005). 'Constitutively high expression of the histidine biosynthetic pathway contributes to nickel tolerance in hyperaccumulator plants.' *The Plant Cell*, 17/7: 2089–106.

Käll, L., Krogh, A., & Sonnhammer, E. L. L. (2004). 'A combined transmembrane topology and signal peptide prediction method.' *Journal of Molecular Biology*, 338/5: 1027–36.

Kampfenkel, K., Kushnir, S., Babiychuk, E., Inzé, D., & Van Montagu, M. (1995). 'Molecular characterization of a putative *Arabidopsis thaliana* copper transporter and its yeast homolog.' *Journal of Biological Chemistry*, 270/47: 28479–86.

Kerkeb, L., & Krämer, U. (2003). 'The role of free histidine in xylem loading of nickel in *Alyssum lesbiacum* and *Brassica juncea*.' *Plant physiology*, 131/2: 716–24.

Kim, S. A., Punshon, T., Lanzirotti, A., Li, L., Alonso, J. M., Ecker, J. R., Kaplan, J., & Guerinot, M.L. (2006). 'Localization of iron in *Arabidopsis* seed requires the vacuolar membrane transporter VIT1.' *Science*, 314/5803: 1295–8.

Krämer, U. (2010). 'Metal hyperaccumulation in plants.' *Annual Review of Plant Biology*, 61: 517–34.

Krämer, U., Talke, I. N., & Hanikenne, M. (2007). 'Transition metal transport.' *FEBS Letters*, 581/12: 2263–72.

Krogh, A., Larsson, B., Von Heijne, G., & Sonnhammer, E. (2001). 'Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.' *Journal of Molecular Biology*, 305/3: 567–80.

Kumar, S., Stecher, G., & Tamura, K. (2016). 'MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets.' *Molecular Biology and Evolution*, 33/7: 1870–4.

Li, J., Park, E., Von Arnim, A. G., & Nebenführ, A. (2009). 'The FAST technique: a simplified *Agrobacterium*-based transformation method for transient gene expression analysis in seedlings of *Arabidopsis* and other plant species.' *Plant Methods*, 5: 1-15.

Li, L., He, Z., Pandey, G. K., Tsuchiya, T., & Luan, S. (2002). 'Functional cloning and characterization of a plant efflux carrier for multidrug and heavy metal detoxification.' *Journal of Biological Chemistry*, 277/7: 5360–8.

Li, L., & Kaplan, J. (1998). 'Defects in the yeast high affinity iron transport system result in increased metal sensitivity because of the increased expression of transporters with a broad transition metal specificity.' *Journal of Biological Chemistry*, 273/35: 22181–7.

Liu, J., Li, Y., Wang, W., Gai, J., & Li, Y. (2016). 'Genome-wide analysis of MATE transporters and expression patterns of a subgroup of *MATE* genes in response to aluminum toxicity in soybean.' *BMC Genomics*, 17/1: 1–15.

Magalhaes, J. V, Liu, J., Guimarães, C. T., Lana, U. G. P., Alves, V. M. C., Wang, Y. H., Schaffert, R. E., Hoekenga, O. A., Piñeros, M. A., Shaff, J. E., Klein, P. E., Carneiro, N. P., Coelho, C. M., Trick, H. N., & Kochian, L. V. (2007). 'A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in

sorghum.' *Nature Genetics*, 39/9: 1156–61.

Mahar, A., Wang, P., Ali, A., Awasthi, M. K., Lahori, A. H., Wang, Q., Li, R., & Zhang, Z. (2016). 'Challenges and opportunities in the phytoremediation of heavy metals contaminated soils: a review.' *Ecotoxicology and Environmental Safety*, 126: 111–21.

Marion, J., Bach, L., Bellec, Y., Meyer, C., Gissot, L., & Faure, J. D. (2008). 'Systematic analysis of protein subcellular localization and interaction using high-throughput transient transformation of *Arabidopsis* seedlings.' *The Plant Journal*, 56/1: 169–79.

Maron, L. G., Piñeros, M. A., Guimarães, C. T., Magalhaes, J. V., Pleiman, J. K., Mao, C., Shaff, J., Belicaus, S. N. J., & Kochian, L. V. (2010). 'Two functionally distinct members of the MATE (multi-drug and toxic compound extrusion) family of transporters potentially underlie two major aluminum tolerance QTLs in maize.' *Plant Journal*, 61/5: 728–40.

Meier, S., Adams, N., Wolf, M., Bankwill, K., Muasya, M., Gehring, C., Bishop, J., & Ingle, R. A. (in review.). 'RNA-Seq analysis of *Senecio coronatus* (Asteraceae) suggests that mis-regulation of iron uptake may underlie the evolution of nickel hyperaccumulation in this species.'

Mengel, K., Kirkby, E. A., Kosegarten, H., & Appel, T. (2001). *Principles of Plant Nutrition*. Springer: Netherlands.

Merlot, S., Hannibal, L., Martins, S., Martinelli, L., Amir, H., Lebrun, M., & Thomine, S. (2014). 'The metal transporter PgIREG1 from the hyperaccumulator *Psychotria gabriellae* is a candidate gene for nickel tolerance and accumulation.' *Journal of Experimental Botany*, 65/6: 1551–64.

Mesjasz-Przybyłowicz, J., & Przybyłowicz, W. J. (2001). 'Phytophagous insects associated with the Ni-hyperaccumulating plant *Berkheya coddii* (Asteraceae) in Mpumalanga, South Africa.' *South African Journal of Science*, 97/11–12: 596–8.

Mesjasz-Przybyłowicz, J., Przybyłowicz, W. J., Prozesky, V. M., & Pineda, C. A. (1997). 'Quantitative micro-PIXE comparison of elemental distribution in Ni-hyperaccumulating and non-accumulating genotypes of *Senecio coronatus*.'

Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 130/1–4: 368–73.

Minguzzi, A., & Vergano, O. (1948). 'Il contenuto di nichel nelle ceneri di *Alyssum bertolonii* Desv.' *Memorie Società Toscana Di Scienze Naturali*, 55: 49–77.

Mizuno, T., Usui, K., Horie, K., Nosaka, S., Mizuno, N., & Obata, H. (2005). 'Cloning of three ZIP/Nramp transporter genes from a Ni hyperaccumulator plant *Thlaspi japonicum* and their Ni2+-transport abilities.' *Plant Physiology and Biochemistry*, 43/8: 793–801.

Morrey, D. R., Balkwill, K., & Balkwill, M. J. (1989). 'Studies on serpentine flora: preliminary analyses of soils and vegetation associated with serpentinite rock formations in the south-eastern Transvaal.' *South African Journal of Botany*, 55/2: 171–7.

Morris, C., Grossl, P. R., & Call, C. A. (2009). 'Elemental allelopathy: processes, progress, and pitfalls.' *Plant Ecology*, 202/1: 1–11.

Morrissey, J., Baxter, I. R., Lee, J., Li, L., Lahner, B., Grotz, N., Kaplan, J., Salt, D. E., & Guerinot, M. L. (2009). 'The ferroportin metal efflux proteins function in iron and cobalt homeostasis in *Arabidopsis*.' *The Plant cell*, 21/10: 3326–38.

Van De Mortel, J. E., Villanueva, L. A., Schat, H., Kwekkeboom, J., Coughlan, S., Moerland, P. D., Ver Loren Van Themaat, E., Koornneef, M., & Aarts, M. G. M. (2006). 'Large expression differences in genes for iron and zinc homeostasis, stress response, and lignin biosynthesis distinguish roots of *Arabidopsis thaliana* and the related metal hyperaccumulator *Thlaspi caerulescens*.' *Plant Physiology*, 142: 1127–47.

Nagajyoti, P. C., Lee, K. D., & Sreekanth, T. V. M. (2010). 'Heavy metals, occurrence and toxicity for plants: a review.' *Environmental Chemistry Letters*, 8/3: 199–216.

Nishida, S., Tsuzuki, C., Kato, A., Aisu, A., Yoshida, J., & Mizuno, T. (2011). 'AtIRT1, the primary iron uptake transporter in the root, mediates excess nickel accumulation in *Arabidopsis thaliana*.' *Plant and Cell Physiology*, 52/8: 1433–42.

Oomen, R. J. F. J., Wu, J., Lelièvre, F., Blanchet, S., Richaud, P., Barbier-Brygoo, H., Aarts, M. G. M., & Thomine, S. (2009). 'Functional characterization of NRAMP3 and NRAMP4 from the metal hyperaccumulator *Thlaspi caerulescens*.' *New*

Phytologist, 181/3: 637–50.

Pence, N. S., Larsen, P. B., Ebbs, S. D., Letham, D. L., Lasat, M. M., Garvin, D. F., Eide, D., & Kochian, L. V. (2000). 'The molecular physiology of heavy metal transport in the Zn/Cd hyperaccumulator *Thlaspi caerulescens*.' *Proceedings of the National Academy of Sciences of the United States of America*, 97/9: 4956–60.

Perea-García, A., Garcia-Molina, A., Andrés-Colás, N., Vera-Sirera, F., Pérez-Amador, M. A., Puig, S., & Peñarrubia, L. (2013). '*Arabidopsis* copper transport protein COPT2 participates in the cross talk between iron deficiency responses and low-phosphate signaling.' *Plant Physiology*, 162/1: 180–94.

Pollard, A. J., Reeves, R. D., & Baker, A. J. M. (2014). 'Facultative hyperaccumulation of heavy metals and metalloids.' *Plant Science*, 217: 8–17.

Puig, S., & Thiele, D. J. (2002). 'Molecular mechanisms of copper uptake and distribution.' *Current Opinion in Chemical Biology*, 6/2: 171–80.

Rascio, N., & Navari-Izzo, F. (2011). 'Heavy metal hyperaccumulating plants: how and why do they do it? And what makes them so interesting?' *Plant Science*, 180/2: 169–81.

Reeves, R., & Baker, A. (2000). 'Metal-accumulating plants.' Raskin I. & Ensley B. (eds) *Phytoremediation of Toxic Metals - Using Plants to Clean Up the Environment*, pp. 193–229. Wiley: New York.

Reeves, R. D., Baker, A. J. M., Borhidi, A., & Berazaín, R. (1996). 'Nickel-accumulating plants from the ancient serpentine soils of Cuba.' *New Phytologist*, 133/2: 217–24.

Reeves, R. D., Brooks, R. R., & Macfarlane, R. M. (1981). 'Nickel Uptake by Californian *Streptanthus* and *Caulanthus* with particular reference to the hyperaccumulator *S. polygaloides* Gray (Brassicaceae).' *American Journal of Botany*, 68/5: 708.

Sagner, S., Kneer, R., Wanner, G., Cosson, J. P., Deus-Neumann, B., & Zenk, M. H. (1998). 'Hyperaccumulation, complexation and distribution of nickel in *Sebertia acuminata*.' *Phytochemistry*, 47/3: 339–47.

Saitou, N., & Nei, M. (1987). 'The neighbor-joining method: a new method for reconstructing phylogenetic trees.' *Molecular Biology and Evolution*, 4/4: 406–

25.

Sancenón, V., Puig, S., Mateu-Andrés, I., Dorcey, E., Thiele, D. J., & Peñarrubia, L. (2004). 'The *Arabidopsis* copper transporter COPT1 functions in root elongation and pollen development.' *Journal of Biological Chemistry*, 279/15: 15348–55.

Schaaf, G., Honsbein, A., Meda, A. R., Kirchner, S., Wipf, D., & Von Wirén, N. (2006). '*AtIREG2* encodes a tonoplast transport protein involved in iron-dependent nickel detoxification in *Arabidopsis thaliana* roots.' *Journal of Biological Chemistry*, 281/35: 25532–40.

Shitan, N., Minami, S., Morita, M., Hayashida, M., Ito, S., Takanashi, K., Omote, H., Moriyama, Y., Sugiyama, A., Goosens, A., Moriyasu, M., & Yazaki, K. (2014). 'Involvement of the leaf-specific multidrug and toxic compound extrusion (MATE) transporter Nt-JAT2 in vacuolar sequestration of nicotine in *Nicotiana tabacum*.' *PLoS ONE*, 9/9: 1–10.

Slavic, K., Krishna, S., Lahree, A., Bouyer, G., Hanson, K. K., Pittman, J. K., Staines, H. M., & Mota, M. M. (2016). 'A vacuolar iron-transporter homologue acts as a detoxifier in *Plasmodium*.' *Nature Communications*, 7: 1–10.

Sobczyk, M. K., Smith, J. A. C., Pollard, A. J., & Filatov, D. A. (2017). 'Evolution of nickel hyperaccumulation and serpentine adaptation in the *Alyssum serpyllifolium* species complex.' *Heredity*, 118/1: 31–41.

Talke, I. N. (2006). 'Zinc-dependent global transcriptional control, transcriptional deregulation, and higher gene copy number for genes in metal homeostasis of the hyperaccumulator *Arabidopsis halleri*.' *Plant Physiology*, 142/1: 148–67.

Tanz, S. K., Castleden, I., Hooper, C. M., Vacher, M., Small, I., & Millar, H. A. (2013). 'SUBA3: a database for integrating experimentation and prediction to define the subcellular location of proteins in *Arabidopsis*.' *Nucleic Acids Research*, 41: 1185–91.

Tauqeer, H. M., Ali, S., Rizwan, M., Ali, Q., Saeed, R., Iftikhar, U., Ahmad, R., Farid, M., & Abassi, G. H. (2016). 'Phytoremediation of heavy metals by *Alternanthera bettzickiana*: growth and physiological response.' *Ecotoxicology and Environmental Safety*, 126: 138–46.

Thalius, J. (1588). *Sylva Hercynica, sive catalogus plantarum sponte nascentium in*

montibus et locis vicinis Hercynae. Feyerabend: Frankfurt.

Ueki, S., Lacroix, B., Krichevsky, A., Lazarowitz, S. G., & Citovsky, V. (2009). 'Functional transient genetic transformation of *Arabidopsis* leaves by biolistic bombardment.' *Nature Protocols*, 4/1: 71–7.

Verbruggen, N., Hermans, C., & Schat, H. (2009). 'Molecular mechanisms of metal hyperaccumulation in plants.' *New Phytologist*, 181/4: 759–76.

Vert, G., Grotz, N., Dédaldéchamp, F., Gaymard, F., Guerinot, M. L., Briat, J. F., & Curie, C. (2002). 'IRT1, an *Arabidopsis* transporter essential for iron uptake from the soil and for plant growth.' *The Plant Cell*, 14/6: 1223–33.

Weber, M., Harada, E., Vess, C., Roepenack-Lahaye, E. V., & Clemens, S. (2004). 'Comparative microarray analysis of *Arabidopsis thaliana* and *Arabidopsis halleri* roots identifies nicotianamine synthase, a ZIP transporter and other genes as potential metal hyperaccumulation factors.' *The Plant Journal*, 37/2: 269–81.

Wei, W., Chai, T., Zhang, Y., Han, L., Xu, J., & Guan, Z. (2009). 'The *Thlaspi caerulescens* NRAMP homologue TcNRAMP3 is capable of divalent cation transport.' *Molecular Biotechnology*, 41/1: 15–21.

Williams, L., & Salt, D. E. (2009). 'The plant ionome coming into focus.' *Current Opinion in Plant Biology*, 12/3: 247–9.

Wink, M. (1997). 'Compartmentation of secondary metabolites and xenobiotics in plant vacuoles.' *Advances in Botanical Research*, 25: 141–69.

Xu, K., Huang, X., Wu, M., Wang, Y., Chang, Y., Liu, K., Zhang, J., Zhang, Y., Zhang, F., Yi, L., Li, T., Wang, R., Tan, G., & Li, C. (2014). 'A rapid , highly efficient and economical method of *Agrobacterium*-mediated *in planta* transient transformation in living onion epidermis.' *PLoS ONE,* 9/1: 1–7.

Yadav, S. K. (2010). 'Heavy metals toxicity in plants: an overview on the role of glutathione and phytochelatins in heavy metal stress tolerance of plants.' *South African Journal of Botany*, 76/2: 167–79.

Ye, G. N., Stone, D., Pang, S. Z., Creely, W., Gonzalez, K., & Hinchee, M. (1999). '*Arabidopsis* ovule is the target for *Agrobacterium in planta* vacuum infiltration transformation.' *Plant Journal*, 19/3: 249–57.

Yesilirmak, F., & Sayers, Z. (2009). 'Heterelogous expression of plant genes.'

International Journal of Plant Genomics, 2009: 1-16.

Yu, C. S., Chen, Y. C., Lu, C. H., & Hwang, J. K. (2008). 'Prediction of protein subcellular localization.' *Proteins*, 70/2: 311–9.

Yuan, M., Li, X., Xiao, J., & Wang, S. (2011). 'Molecular and functional analyses of COPT/Ctr-type copper transporter-like gene family in rice.' *BMC Plant Biology*, 11/1: 1–12.

Zhang, G., Lu, S., Chen, T. A., Funk, C. R., & Meyer, W. A. (2003). 'Transformation of triploid bermudagrass (*Cynodon dactylon* x *C. transvaalensis* cv. TifEagle) by means of biolistic bombardment.' *Plant Cell Reports*, 21/9: 860–4.

Zhang, L., Angle, J. S., & Chaney, R. L. (2007). 'Do high-nickel leaves shed by the nickel hyperaccumulator *Alyssum murale* inhibit seed germination of competing plants?' *New Phytologist*, 173/3: 509–16.

Zhang, Y., Xu, Y. H., Yi, H. Y., & Gong, J. M. (2012). 'Vacuolar membrane transporters OsVIT1 and OsVIT2 modulate iron translocation between flag leaves and seeds in rice.' *Plant Journal*, 72/3: 400–10.

Zhao, H., & Eide, D. (1996). 'The yeast *ZRT1* gene encodes the zinc transporter protein of a high-affinity uptake system induced by zinc limitation.' *Proceedings of the National Academy of Sciences of the United States of America*, 93/6: 2454–8.

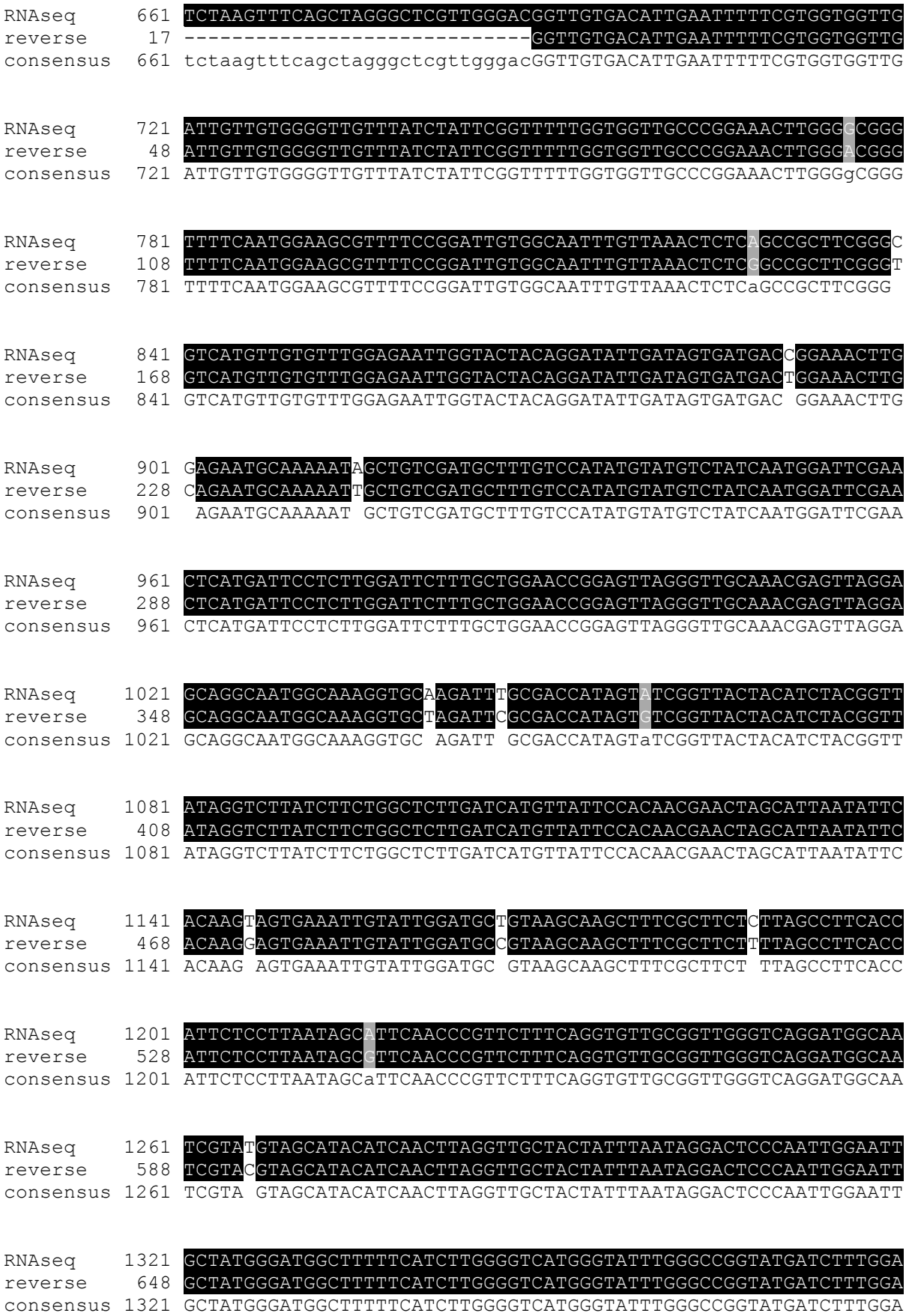
Zhu, W., Zuo, R., Zhou, R., Huang, J., Tang, M., Cheng, X., Liu, Y., Tong, C., Xiang, Y., Dong, C., & Liu, S. (2016). 'Vacuolar Iron Transporter BnMEB2 is involved in enhancing iron tolerance of *Brassica napus*.' *Frontiers in Plant Science*, 7: 1-13.

Zuckerkandl, E., & Pauling, L. (1965). 'Evolutionary divergence and convergence in proteins.' V. Bryson and H.J. Vogel (eds.) *Evolving Genes and Proteins* , pp. 97–166. Academic Press: New York.

VI APPENDIX 1

MATE(a) – RNA-Seq derived sequence ("RNAseq") and forward/reverse cDNA sequences

```
RNAseq       1 ATGGCCATCAATGGCACTGGTGTTCGAGTAGACGAACCAATACTAGTACCTTTGTTGGAT
forward      1 ATGGCCATCAATGGCACTGGTGTTCGAGTAGACGAACCAATACTAGTACCTTTGTTGGAT
consensus    1 ATGGCCATCAATGGCACTGGTGTTCGAGTAGACGAACCAATACTAGTACCTTTGTTGGAT

RNAseq      61 CATTCTTCATCAAACGTCGACTTGGTTTCGAAAACAACTGACCAAAAACTCGATCTCTCG
forward     61 CATTCTTCATCAAACGTCGACTTGGTTTCGAAAACAACTGACCAAAAACTCGATCTCTCG
consensus   61 CATTCTTCATCAAACGTCGACTTGGTTTCGAAAACAACTGACCAAAAACTCGATCTCTCG

RNAseq     121 TTTCGACAACGATACTTGATCGAATCAAAGAAGCTATGGCATATAGTTGGTCCTGCTATC
forward    121 TTTCGACAACGATACTTGATCGAATCAAAGAAGCTATGGCATATAGTTGGTCCTGCTATC
consensus  121 TTTCGACAACGATACTTGATCGAATCAAAGAAGCTATGGCATATAGTTGGTCCTGCTATC

RNAseq     181 TTCAGCCGCATAGCTTCCTACTCCATGTTTGTTATCACCCAAGCTTTTGCCGGACATCTT
forward    181 TTCAGCCGCATAGCTTCCTACTCCATGTTTGTTATCACCCAAGCTTTTGCCGGACATCTT
consensus  181 TTCAGCCGCATAGCTTCCTACTCCATGTTTGTTATCACCCAAGCTTTTGCCGGACATCTT

RNAseq     241 GGTGATCTTGAACTAGCTGGTATTTCCATTGCTACAAGTGTCATTGTTGGCTTCGACTTT
forward    241 GGTGATCTTGAACTAGCTGGTATTTCCATTGCTACAAGTGTCATTGTTGGCTTCGACTTT
consensus  241 GGTGATCTTGAACTAGCTGGTATTTCCATTGCTACAAGTGTCATTGTTGGCTTCGACTTT

RNAseq     301 GGCCTCTTGTTAGGGATGGCAAGTGCATTAGAAACATTGTGTGGTCAAGCATACGGAGCG
forward    301 GGCCTCTTGTTAGGGATGGCAAGTGCATTAGAAACATTGTGTGGTCAAGCATACGGAGCG
consensus  301 GGCCTCTTGTTAGGGATGGCAAGTGCATTAGAAACATTGTGTGGTCAAGCATACGGAGCG

RNAseq     361 AAAAACTATCGAATGTTGGGAGTATATTTACAACGCTCATGGATCGTGCTTTTCGTATGT
forward    361 AAAAACTATCGAATGTTGGGAGTATATTTACAACGCTCATGGATCGTGCTTTTCGTATGT
consensus  361 AAAAACTATCGAATGTTGGGAGTATATTTACAACGCTCATGGATCGTGCTTTTCGTATGT

RNAseq     421 TGTGTACTACTTTTACCCTTATACATCTTCGCTACACCGGTACTAAAACTCCTCGGACAA
forward    421 TGTGTACTACTTTTACCCTTATACATCTTCGCTACACCGGTACTAAAACTCCTCGGACAA
consensus  421 TGTGTACTACTTTTACCCTTATACATCTTCGCTACACCGGTACTAAAACTCCTCGGACAA

RNAseq     481 CCAGCAGATATAGCTGAGCTCTCTGGCATAGTGTCAATGAGTTTAATACCACTCCATTTC
forward    481 CCAGCAGATATAGCTGAGCTCTCTGGCATAGTGT--------------------------
consensus  481 CCAGCAGATATAGCTGAGCTCTCTGGCATAGTGTcaatgagtttaataccactccatttc

RNAseq     541 AGCCTATGTTTTCAATTCCCATTACAAAGGTTCCTTCAAAGCCAGCTTAAAACATTCGTA
forward    515 ------------------------------------------------------------
consensus  541 agcctatgttttcaattcccattacaaaggttccttcaaagccagcttaaaacattcgta

RNAseq     601 ATTGCGTGGGTTTCATTGGGAGCGTTGGTGGTTCATTTGTTCATGAGTTGGCTTGTTGTG
forward    515 ------------------------------------------------------------
consensus  601 attgcgtgggtttcattgggagcgttggtggttcatttgttcatgagttggcttgttgtg
```


RNAseq 661 TCTAAGTTTCAGCTAGGGCTCGTTGGGACGGTTGTGACATTGAATTTTTCGTGGTGGTTG
reverse 17 ----------GGTTGTGACATTGAATTTTTCGTGGTGGTTG
consensus 661 tctaagtttcagctagggctcgttgggacGGTTGTGACATTGAATTTTTCGTGGTGGTTG
RNAseq 721 ATTGTTGTGGGGTTGTTTATCTATTCGGTTTTTGGTGGTTGCCCGGAAACTTGGGGCGGG
reverse 48 ATTGTTGTGGGGTTGTTTATCTATTCGGTTTTTGGTGGTTGCCCGGAAACTTGGGACGGG
consensus 721 ATTGTTGTGGGGTTGTTTATCTATTCGGTTTTTGGTGGTTGCCCGGAAACTTGGGgCGGG
RNAseq 781 TTTTCAATGGAAGCGTTTTCCGGATTGTGGCAATTTGTTAAACTCTCAGCCGCTTCGGGC
reverse 108 TTTTCAATGGAAGCGTTTTCCGGATTGTGGCAATTTGTTAAACTCTCGGCCGCTTCGGGT
consensus 781 TTTTCAATGGAAGCGTTTTCCGGATTGTGGCAATTTGTTAAACTCTCaGCCGCTTCGGG
RNAseq 841 GTCATGTTGTGTTTGGAGAATTGGTACTACAGGATATTGATAGTGATGACCGGAAACTTG
reverse 168 GTCATGTTGTGTTTGGAGAATTGGTACTACAGGATATTGATAGTGATGACTGGAAACTTG
consensus 841 GTCATGTTGTGTTTGGAGAATTGGTACTACAGGATATTGATAGTGATGAC GGAAACTTG
RNAseq 901 GAGAATGCAAAAATAGCTGTCGATGCTTTGTCCATATGTATGTCTATCAATGGATTCGAA
reverse 228 CAGAATGCAAAAATTGCTGTCGATGCTTTGTCCATATGTATGTCTATCAATGGATTCGAA
consensus 901 AGAATGCAAAAAT GCTGTCGATGCTTTGTCCATATGTATGTCTATCAATGGATTCGAA
RNAseq 961 CTCATGATTCCTCTTGGATTCTTTGCTGGAACCGGAGTTAGGGTTGCAAACGAGTTAGGA
reverse 288 CTCATGATTCCTCTTGGATTCTTTGCTGGAACCGGAGTTAGGGTTGCAAACGAGTTAGGA
consensus 961 CTCATGATTCCTCTTGGATTCTTTGCTGGAACCGGAGTTAGGGTTGCAAACGAGTTAGGA
RNAseq 1021 GCAGGCAATGGCAAAGGTGCAAGATTTGCGACCATAGTATCGGTTACTACATCTACGGTT
reverse 348 GCAGGCAATGGCAAAGGTGCTAGATTCGCGACCATAGTGTCGGTTACTACATCTACGGTT
consensus 1021 GCAGGCAATGGCAAAGGTGC AGATT GCGACCATAGTaTCGGTTACTACATCTACGGTT
RNAseq 1081 ATAGGTCTTATCTTCTGGCTCTTGATCATGTTATTCCACAACGAACTAGCATTAATATTC
reverse 408 ATAGGTCTTATCTTCTGGCTCTTGATCATGTTATTCCACAACGAACTAGCATTAATATTC
consensus 1081 ATAGGTCTTATCTTCTGGCTCTTGATCATGTTATTCCACAACGAACTAGCATTAATATTC
RNAseq 1141 ACAAGTAGTGAAATTGTATTGGATGCTGTAAGCAAGCTTTCGCTTCTCTTAGCCTTCACC
reverse 468 ACAAGGAGTGAAATTGTATTGGATGCCGTAAGCAAGCTTTCGCTTCTTTTAGCCTTCACC
consensus 1141 ACAAG AGTGAAATTGTATTGGATGC GTAAGCAAGCTTTCGCTTCT TTAGCCTTCACC
RNAseq 1201 ATTCTCCTTAATAGCATTCAACCCGTTCTTTCAGGTGTTGCGGTTGGGTCAGGATGGCAA
reverse 528 ATTCTCCTTAATAGCGTTCAACCCGTTCTTTCAGGTGTTGCGGTTGGGTCAGGATGGCAA
consensus 1201 ATTCTCCTTAATAGCaTTCAACCCGTTCTTTCAGGTGTTGCGGTTGGGTCAGGATGGCAA
RNAseq 1261 TCGTATGTAGCATACATCAACTTAGGTTGCTACTATTTAATAGGACTCCCAATTGGAATT
reverse 588 TCGTACGTAGCATACATCAACTTAGGTTGCTACTATTTAATAGGACTCCCAATTGGAATT
consensus 1261 TCGTA GTAGCATACATCAACTTAGGTTGCTACTATTTAATAGGACTCCCAATTGGAATT
RNAseq 1321 GCTATGGGATGGCTTTTTCATCTTGGGGTCATGGGTATTTGGGCCGGTATGATCTTTGGA
reverse 648 GCTATGGGATGGCTTTTTCATCTTGGGGTCATGGGTATTTGGGCCGGTATGATCTTTGGA
consensus 1321 GCTATGGGATGGCTTTTTCATCTTGGGGTCATGGGTATTTGGGCCGGTATGATCTTTGGA

```
RNAseq    1381 GGAACTGCATTTCAAACCGTGGTATTGGCTATAATCACAAGTCGTTGCGATTGGGAAAAA
reverse    708 GGAACTGCATTTCAAACCGTGGTATTGGCTATAATCACAAGTCGTTGCGATTGGGAAAAA
consensus 1381 GGAACTGCATTTCAAACCGTGGTATTGGCTATAATCACAAGTCGTTGCGATTGGGAAAAA

RNAseq    1441 GAGGCTCTAAGAGCAAGTACACATGTAAAGAAATGGGCAGTTGTTCATTGA
reverse    768 GAGGCTCTAAGAGCAAGTACACATGTAAAGAAATGGGCAGTTGTTCATTGA
consensus 1441 GAGGCTCTAAGAGCAAGTACACATGTAAAGAAATGGGCAGTTGTTCATTGA
```

MATE(b) RNA-Seq derived sequence ("RNAseq") and three reverse cDNA sequences

```
RNAseq           1 TAGGGCTCGTTGGGACGGTTGTGACATTGAATTTTTCGTGGTGGTTGATTGTTGTGGGGT
reverse1_r/c_    1 TAGGACTCGTTGGGACGGTTGTGACATTGAATTTTTCGTGGTGGTTGATTGTTGTGGGGT
reverse2_r/c_    1 TAGGACTCGTTGGGACGGTTGTGACATTGAATTTTTCGTGGTGGTTGATTGTTGTGGGGT
reverse3_r/c_    1 TAGGACTCGTTGGGACGGTTGTGACATTGAATTTTTCGTGGTGGTTGATTGTTGTGGGGT
consensus        1 TAGGaCTCGTTGGGACGGTTGTGACATTGAATTTTTCGTGGTGGTTGATTGTTGTGGGGT

RNAseq          61 TGTTTATCTATTCGGTTTTTGGTGGTTGCCCGGAAACTTGGGGCGGGTTTTCAATGGAAG
reverse1_r/c_   61 TGTTTATCTATTCGGTTTTTGGTGGTTGCCCGGAAACTTGGGACGGGTTTTCAATGGAAG
reverse2_r/c_   61 TGTTTATCTATTCGGTTTTTGGTGGTTGCCCGGAAACTTGGGACGGGTTTTCAATGGAAG
reverse3_r/c_   61 TGTTTATCTATTCGGTTTTTGGTGGTTGCCCGGAAACTTGGGACGGGTTTTCAATGGAAG
consensus       61 TGTTTATCTATTCGGTTTTTGGTGGTTGCCCGGAAACTTGGGaCGGGTTTTCAATGGAAG

RNAseq         121 CGTTTTCCGGATTGTGGCAATTTGTTAAACTCTCAGCCGCTTCGGGCGTCATGTTGTGTT
reverse1_r/c_  121 CGTTTTCCGGATTGTGGCAATTTGTTAAACTCTCGGCCGCTTCGGGTGTCATGTTGTGTT
reverse2_r/c_  121 CGTTTTCCGGATTGTGGCAATTTGTTAAACTCTCGGCCGCTTCGGGTGTCATGTTGTGTT
reverse3_r/c_  121 CGTTTTCCGGATTGTGGCAATTTGTTAAACTCTCGGCCGCTTCGGGTGTCATGTTGTGTT
consensus      121 CGTTTTCCGGATTGTGGCAATTTGTTAAACTCTCgGCCGCTTCGGGtGTCATGTTGTGTT

RNAseq         181 TGGAGAATTGGTACTACAGGATATTGATAGTGATGACCGGAAACTTGGAGAATGCAAAAA
reverse1_r/c_  181 TGGAGAATTGGTACTACAGGATATTGATAGTGATGACTGGAAACTTGCAGAATGCAAAAA
reverse2_r/c_  181 TGGAGAATTGGTACTACAGGATATTGATAGTGATGACTGGAAACTTGCAGAATGCAAAAA
reverse3_r/c_  181 TGGAGAATTGGTACTACAGGATATTGATAGTGATGACTGGAAACTTGCAGAATGCAAAAA
consensus      181 TGGAGAATTGGTACTACAGGATATTGATAGTGATGACtGGAAACTTGcAGAATGCAAAAA

RNAseq         241 TAGCTGTCGATGCTTTGTCCATATGTATGTCTATCAATGGATTCGAACTCATGATTCCTC
reverse1_r/c_  241 TTGCTGTCGATGCTTTGTCCATATGTATGTCTATCAATGGATTCGAACTCATGATTCCTC
reverse2_r/c_  241 TTGCTGTCGATGCTTTGTCCATATGTATGTCTATCAATGGATTCGAACTCATGATTCCTC
reverse3_r/c_  241 TTGCTGTCGATGCTTTGTCCATATGTATGTCTATCAATGGATTCGAACTCATGATTCCTC
consensus      241 TtGCTGTCGATGCTTTGTCCATATGTATGTCTATCAATGGATTCGAACTCATGATTCCTC

RNAseq         301 TTGGATTCTTTGCTGGAACCGGAGTTAGGGTTGCAAACGAGTTAGGAGCAGGCAATGGCA
reverse1_r/c_  301 TTGGATTCTTTGCTGGAACCGGAGTTAGGGTTGCAAACGAGTTAGGAGCAGGCAATGGCA
reverse2_r/c_  301 TTGGATTCTTTGCTGGAACCGGAGTTAGGGTTGCAAACGAGTTAGGAGCAGGCAATGGCA
reverse3_r/c_  301 TTGGATTCTTTGCTGGAACCGGAGTTAGGGTTGCAAACGAGTTAGGAGCAGGCAATGGCA
consensus      301 TTGGATTCTTTGCTGGAACCGGAGTTAGGGTTGCAAACGAGTTAGGAGCAGGCAATGGCA

RNAseq         361 AAGGTGCAAGATTTGCGACCATAGTATCGGTTACTACATCTACGGTTATAGGTCTTATCT
reverse1_r/c_  361 AAGGTGCTAGATTCGCGACCATAGTGTCGGTTACTACATCTACGGTTATAGGTCTTATCT
reverse2_r/c_  361 AAGGTGCTAGATTCGCGACCATAGTGTCGGTTACTACATCTACGGTTATAGGTCTTATCT
reverse3_r/c_  361 AAGGTGCTAGATTCGCGACCATAGTGTCGGTTACTACATCTACGGTTATAGGTCTTATCT
consensus      361 AAGGTGCtAGATTcGCGACCATAGTgTCGGTTACTACATCTACGGTTATAGGTCTTATCT

RNAseq         421 TCTGGCTCTTGATCATGTTATTCCACAACGAACTAGCATTAATATTCACAAGTAGTGAAA
reverse1_r/c_  421 TCTGGCTCTTGATCATGTTATTCCACAACGAACTAGCATTAATATTCACAAGGAGTGAAA
reverse2_r/c_  421 TCTGGCTCTTGATCATGTTATTCCACAACGAACTAGCATTAATATTCACAAGGAGTGAAA
reverse3_r/c_  421 TCTGGCTCTTGATCATGTTATTCCACAACGAACTAGCATTAATATTCACAAGGAGTGAAA
consensus      421 TCTGGCTCTTGATCATGTTATTCCACAACGAACTAGCATTAATATTCACAAGgAGTGAAA
```

```
RNAseq          481 TTGTATTGGATGCTGTAAGCAAGCTTTCGCTTCTCTTAGCCTTCACCATTCTCCTTAATA
reverse1_r/c_   481 TTGTATTGGATGCCGTAAGCAAGCTTTCGCTTCTTTTAGCCTTCACCATTCTCCTTAATA
reverse2_r/c_   481 TTGTATTGGATGCCGTAAGCAAGCTTTCGCTTCTTTTAGCCTTCACCATTCTCCTTAATA
reverse3_r/c_   481 TTGTATTGGATGCCGTAAGCAAGCTTTCGCTTCTTTTAGCCTTCACCATTCTCCTTAATA
consensus       481 TTGTATTGGATGCcGTAAGCAAGCTTTCGCTTCTtTTAGCCTTCACCATTCTCCTTAATA

RNAseq          541 GCATTCAACCCGTTCTTTCAGGTGTTGCGGTTGGGTCAGGATGGCAATCGTATGTAGCAT
reverse1_r/c_   541 GCGTTCAACCCGTTCTTTCAGGTGTTGCGGTTGGGTCAGGATGGCAATCGTACGTAGCAT
reverse2_r/c_   541 GCGTTCAACCCGTTCTTTCAGGTGTTGCGGTTGGGTCAGGATGGCAATCGTATGTAGCAT
reverse3_r/c_   541 GCGTTCAACCCGTTCTTTCAGGTGTTGCGGTTGGGTCAGGATGGCAATCGTACGTAGCAT
consensus       541 GCgTTCAACCCGTTCTTTCAGGTGTTGCGGTTGGGTCAGGATGGCAATCGTA GTAGCAT

RNAseq          601 ACATCAACTTAGGTTGCTACTATTTAATAGGACTCCCAATTGGAATTGCTATGGGATGGC
reverse1_r/c_   601 ACATCAACTTAGGTTGCTACTATTTAATAGGACTCCCAATTGGAATTGCTATGGGATGGC
reverse2_r/c_   601 ACATCAACTTAGGTTGCTACTATTTAATAGGACTCCCAATTGGAATTGCTATGGGATGGC
reverse3_r/c_   601 ACATCAACTTAGGTTGCTACTATTTAATAGGACTCCCAATTGGAATTGCTATGGGATGGC
consensus       601 ACATCAACTTAGGTTGCTACTATTTAATAGGACTCCCAATTGGAATTGCTATGGGATGGC

RNAseq          661 TTTTTCATCTTGGGGTCATGGGTATTTGGGCCGGTATGATCTTTGGAGGAACTGCATTTC
reverse1_r/c_   661 TTTTTCATCTTGGGGTCATGGGTATTTGGGCCGGTATGATCTTTGGAGGAACTGCATTTC
reverse2_r/c_   661 TTTTTCATCTTGGGGTCATGGGTATTTGGGCCGGTATGATCTTTGGAGGAACTGCATTTC
reverse3_r/c_   661 TTTTTCATCTTGGGGTCATGGGTATTTGGGCCGGTATGATCTTTGGAGGAACTGCATTTC
consensus       661 TTTTTCATCTTGGGGTCATGGGTATTTGGGCCGGTATGATCTTTGGAGGAACTGCATTTC

RNAseq          721 AAACCGTGGTATTGGCTATAATCACAAGTCGTTGCGATTGGGAAAAAGAGGCTCTAAGAG
reverse1_r/c_   721 AAACCGTGGTATTGGCTATAATCACAAGTCGTTGCGATTGGGAAAAAGAGGCTCTAAGAG
reverse2_r/c_   721 AAACCGTGGTATTGGCTATAATCACAAGTCGTTGCGATTGGGAAAAAGAGGCTCTAAGAG
reverse3_r/c_   721 AAACCGTGGTATTGGCTATAATCACAAGTCGTTGCGATTGGGAAAAAGAGGCTCTAAGAG
consensus       721 AAACCGTGGTATTGGCTATAATCACAAGTCGTTGCGATTGGGAAAAAGAGGCTCTAAGAG

RNAseq          781 CAAGTACACATGTAAAGAAATGGGCAGTTGTTCATTGA
reverse1_r/c_   781 CAAGTACACATGTAAAGAAATGGGCAGTTGTTCATTGA
reverse2_r/c_   781 CAAGTACACATGTAAAGAAATGGGCAGTTGTTCATTGA
reverse3_r/c_   781 CAAGTACACATGTAAAGAAATGGGCAGTTGTTCATTGA
consensus       781 CAAGTACACATGTAAAGAAATGGGCAGTTGTTCATTGA
```

VIT - RNA-Seq derived sequence ("RNAseq") and forward/reverse cDNA sequences

```
RNAseq       1 ATGGCCGGAAATATTGGTGGCGCCACCATAACAACAACCCCTTTTCTCACAAACGGTGCG
forward      1 ATGGCCGGAAATATTGGTGGCGCCACCATAACAACAACCCCTTTTCTCACAAACGGTGCG
consensus    1 ATGGCCGGAAATATTGGTGGCGCCACCATAACAACAACCCCTTTTCTCACAAACGGTGCG

RNAseq      61 GCGGCGGAGGAGAATGGAAGTGGCAGAGAAAGACCAAAAGAACCATGGAAAGGAGAACTA
forward     61 GCGGCGGAGGAGAATGGAAGTGGCAGAGAAAGACCAAAAGAACCATGGAAAGGAGAACTA
consensus   61 GCGGCGGAGGAGAATGGAAGTGGCAGAGAAAGACCAAAAGAACCATGGAAAGGAGAACTA

RNAseq     121 GTTAAGAGCATTGTTTATGCCGGACTTGACGCCATCGTCACCTCCTTTTCTCTCATTTCT
forward    121 GTTAAGAGCATTGTTTATGCCGGACTTGACGCCATCGTCACCTCCTTTTCTCTCATTTCT
consensus  121 GTTAAGAGCATTGTTTATGCCGGACTTGACGCCATCGTCACCTCCTTTTCTCTCATTTCT

RNAseq     181 TCTATCTCGGCTGGCCGTTTATCTTCCGTTGATGTATTGGTGTTGGGGTTTGCTAATTTG
forward    181 TCTATCTCGGCTGGCCGTTTATCTTCCGTTGATGTATTGGTGTTGGGGTTTGCTAATTTG
reverse      4 ------------------------TCCGTTGATGTATTGGTGTTGGGGTTTGCTAATTTG
consensus  181 TCTATCTCGGCTGGCCGTTTATCTTCCGTTGATGTATTGGTGTTGGGGTTTGCTAATTTG

RNAseq     241 GTGGCTGATGGGATATCAATGGGGTTTGGAGACTATGTTTCTAGTAACACCGAGCGAGAT
forward    241 GTGGCTGATGGGATATCAATGGGGTTTGGAGACTATGTTTCTAGTAACACCGAGCGAGAT
reverse     40 GTGGCTGATGGGATATCAATGGGGTTTGGAGACTATGTTTCTAGTAACACCGAGCGAGAT
consensus  241 GTGGCTGATGGGATATCAATGGGGTTTGGAGACTATGTTTCTAGTAACACCGAGCGAGAT

RNAseq     301 GTGGCGGCCAAGGAGCGGCTTGTGACAGAATGGGAGGTAGCTAACCGTCGGAGGAACCAA
forward    301 GTGGCGGCCAAGGAGCGGCTTGTGACAGAATGGGAGGTAGCTAACCGTCGGAGGAACCAA
reverse    100 GTGGCGGCCAAGGAGCGGCTTGTGACAGAATGGGAGGTAGCTAACCGTCGGAGGAACCAA
consensus  301 GTGGCGGCCAAGGAGCGGCTTGTGACAGAATGGGAGGTAGCTAACCGTCGGAGGAACCAA

RNAseq     361 GAACAAGAATTACTTGACCGGTATCAAGATCTTGGCATGAACATCCAAGATGCAACCACG
forward    361 GAACAAGAATTACTTGACCGGTATCAAGATCTTGGCATGAACATCCAAGATGCAACCACG
reverse    160 GAACAAGAATTACTTGACCGGTATCAAGATCTTGGCATGAACATCCAAGATGCAACCACG
consensus  361 GAACAAGAATTACTTGACCGGTATCAAGATCTTGGCATGAACATCCAAGATGCAACCACG

RNAseq     421 GTTGTGAGCATCTTCGCAAAGTACGGAGATATAATGGTGGACGAAAAGATGATCCACAAA
forward    421 GTTGTGAGCATCTTCGCAAAGTACGGAGATATAATGGTGGACGAAAAGATGATCCACAAA
reverse    220 GTTGTGAGCATCTTCGCAAAGTACGGAGATATAATGGTGGACGAAAAGATGATCCACAAA
consensus  421 GTTGTGAGCATCTTCGCAAAGTACGGAGATATAATGGTGGACGAAAAGATGATCCACAAA

RNAseq     481 GGAACCTTATCTCCAGATGATGGTGAGAAGCCATGGAAGAAAGGTCTCATAACATTTGTA
forward    481 GGAACCTTATCTCCAGATGATGGT----------------------------------
reverse    280 GGAACCTTATCTCCAGATGATGGTGAGAAGCCATGGAAGAAAGGTCTCATAACATTTGTA
consensus  481 GGAACCTTATCTCCAGATGATGGTgagaagccatggaagaaaggtctcataacatttgta

RNAseq     541 GCCTTTTTAGTTTTCGGTAGTGCACCAATCCTTGCATTCATCATTCTCATCCCATTCACA
reverse    340 GCCTTTTTAGTTTTCGGTAGTGCACCAATCCTTGCATTCATCATTCTCATCCCATTCACA
consensus  541 GCCTTTTTAGTTTTCGGTAGTGCACCAATCCTTGCATTCATCATTCTCATCCCATTCACA

RNAseq     601 CACAACGACACTCACAAGTTCATAGGAGCATGTATTCTCTCCGCCCTTGCCCTGGCGGCA
reverse    400 CACAACGACACTCACAAGTTCATAGGAGCATGTATTCTCTCCGCCCTTGCCCTGGCGGCA
consensus  601 CACAACGACACTCACAAGTTCATAGGAGCATGTATTCTCTCCGCCCTTGCCCTGGCGGCA
```

```
RNAseq      661 CTTGGAATCGCCAAGGCAAAGATAGCCGGACAGAATCACATGTTGTCCGCCGGTGGGACC
reverse     460 CTTGGAATCGCCAAGGCAAAGATAGCCGGACAGAATCACATGTTGTCCGCCGGTGGGACC
consensus   661 CTTGGAATCGCCAAGGCAAAGATAGCCGGACAGAATCACATGTTGTCCGCCGGTGGGACC

RNAseq      721 TTGTTCAATGGTGCACTTGCTGGTTTTGCAGCTTATGCAATTGGTTGGGTGCTAAGGGAT
reverse     520 TTGTTCAATGGTGCACTTGCTGGTTTTGCAGCTTATGCAATTGGTTGGGTGCTAAGGGAT
consensus   721 TTGTTCAATGGTGCACTTGCTGGTTTTGCAGCTTATGCAATTGGTTGGGTGCTAAGGGAT

RNAseq      781 GTGGCTGGTTTGGAAGATTAG
reverse     580 GTGGCTGGTTTGGAAGATTAG
consensus   781 GTGGCTGGTTTGGAAGATTAG
```

COP - RNA-Seq derived sequence ("RNAseq") and forward cDNA sequence

```
RNAseq       1 ATGATGCATATGACTTTTTGTTGGGGCACCAACGTTACTCTACTAATCGATTCATGGAAA
forward      1 ATGATGCATATGACTTTTTGTTGGGGCACCAACGTTACTCTACTAATCGATTCATGGAAA
consensus    1 ATGATGCATATGACTTTTTGTTGGGGCACCAACGTTACTCTACTAATCGATTCATGGAAA

RNAseq      61 ACGGACTCATGGTTTAGTTACTCCCTCGCGTTAATCATTTGTTTCATATTTTCTGCCTTC
forward     61 ACGGACTCATGGTTTAGTTACTCCCTCGCGTTAATCATTTGTTTCATATTTTCTGCCTTC
consensus   61 ACGGACTCATGGTTTAGTTACTCCCTCGCGTTAATCATTTGTTTCATATTTTCTGCCTTC

RNAseq     121 TACCAGTTTATGGAAGATATTCGTCTCCGATTTAAGCTCCTTTCGTCTTCAGTCACTGCT
forward    121 TACCAGTTTATGGAAGATATTCGTCTCCGATTTAAGCTCCTTTCGTCTTCAGTCACCGCT
consensus  121 TACCAGTTTATGGAAGATATTCGTCTCCGATTTAAGCTCCTTTCGTCTTCAGTCAC GCT

RNAseq     181 GTTGGAGCGGTTGAGAATGCTCCTCTAATTTACAACAAGTTTTTCTCTGGTGGAAGACGA
forward    181 GTTGGAGCGGTTGAGAATGCTCCTCTAATTTACAACAAGTTTTTCTCTGGTGGAAGACGA
consensus  181 GTTGGAGCGGTTGAGAATGCTCCTCTAATTTACAACAAGTTTTTCTCTGGTGGAAGACGA

RNAseq     241 GCGAGGTTTGCTGGATCCTTACTGTTTGGGATAAACTCTGGCCTCAATTACTTTTTGATG
forward    241 GCGAGGTTTGCTGGATCCTTACTGTTTGGGATAAACTCTGGCCTCAATTACTTTTTGATG
consensus  241 GCGAGGTTTGCTGGATCCTTACTGTTTGGGATAAACTCTGGCCTCAATTACTTTTTGATG

RNAseq     301 CTTGCTGTGATGTCGTTTAACGTTGGTGTGTTTGTGGTGATTGTGGCTGGACTAGCGGTC
forward    301 CTTGCTGTGATGTCGTTTAACGTTGGTGTGTTTGTGGTGATTGTGGCTGGACTAGCGGTC
consensus  301 CTTGCTGTGATGTCGTTTAACGTTGGTGTGTTTGTGGTGATTGTGGCTGGACTAGCGGTC

RNAseq     361 GGGTACTGGTTGTTTAGGAGTGCTGATGATGAACAGATTACGCTTCTTGATGA
forward    361 GGGTACTGGTTGTTTAGGAGTGCTGATGATGAACAGATTACGCTTCTTGATGA
consensus  361 GGGTACTGGTTGTTTAGGAGTGCTGATGATGAACAGATTACGCTTCTTGATGA
```

VII APPENDIX 2

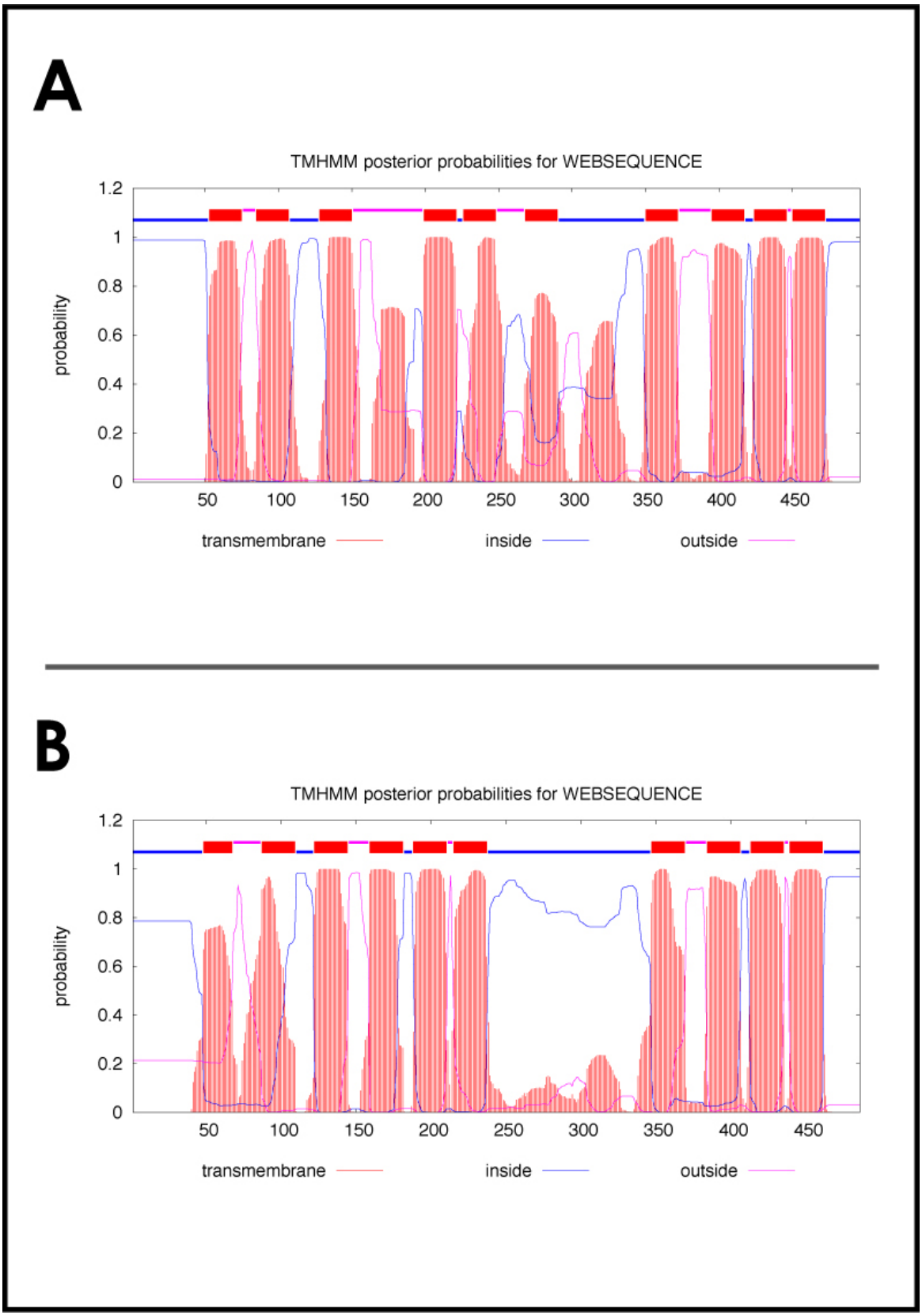


Figure A2(1). Transmembrane domain predictions for ScMATE (**A**) and AtMATE (**B**) generated using THHMM. The total number of TMDs predicted is ten.

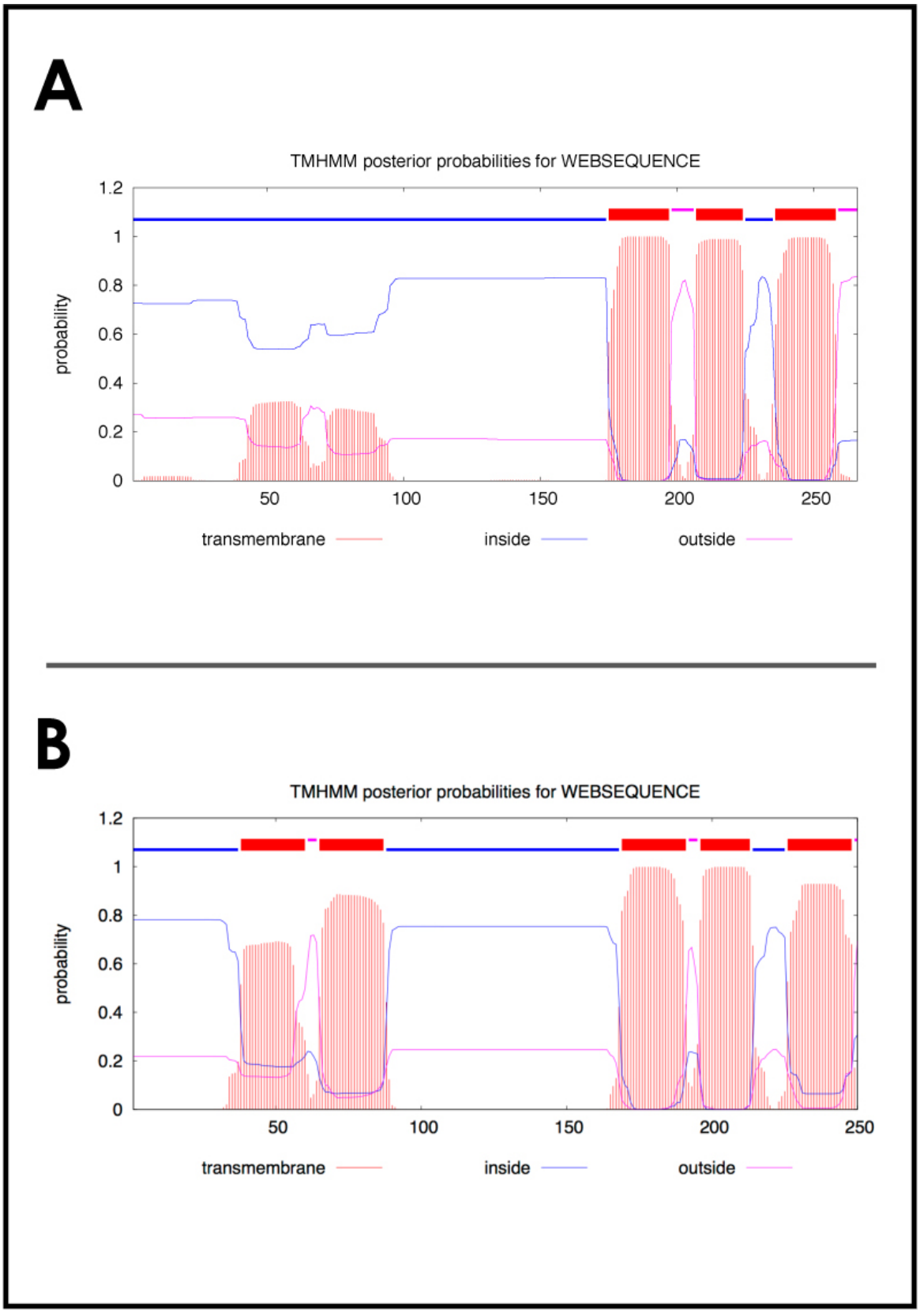


Figure A2(2). Transmembrane domain predictions for ScVIT (**A**) and AtVIT (**B**) generated using TMHMM. The total number of TMDs predicted ScVIT is and AtVIT is three and five respectively.

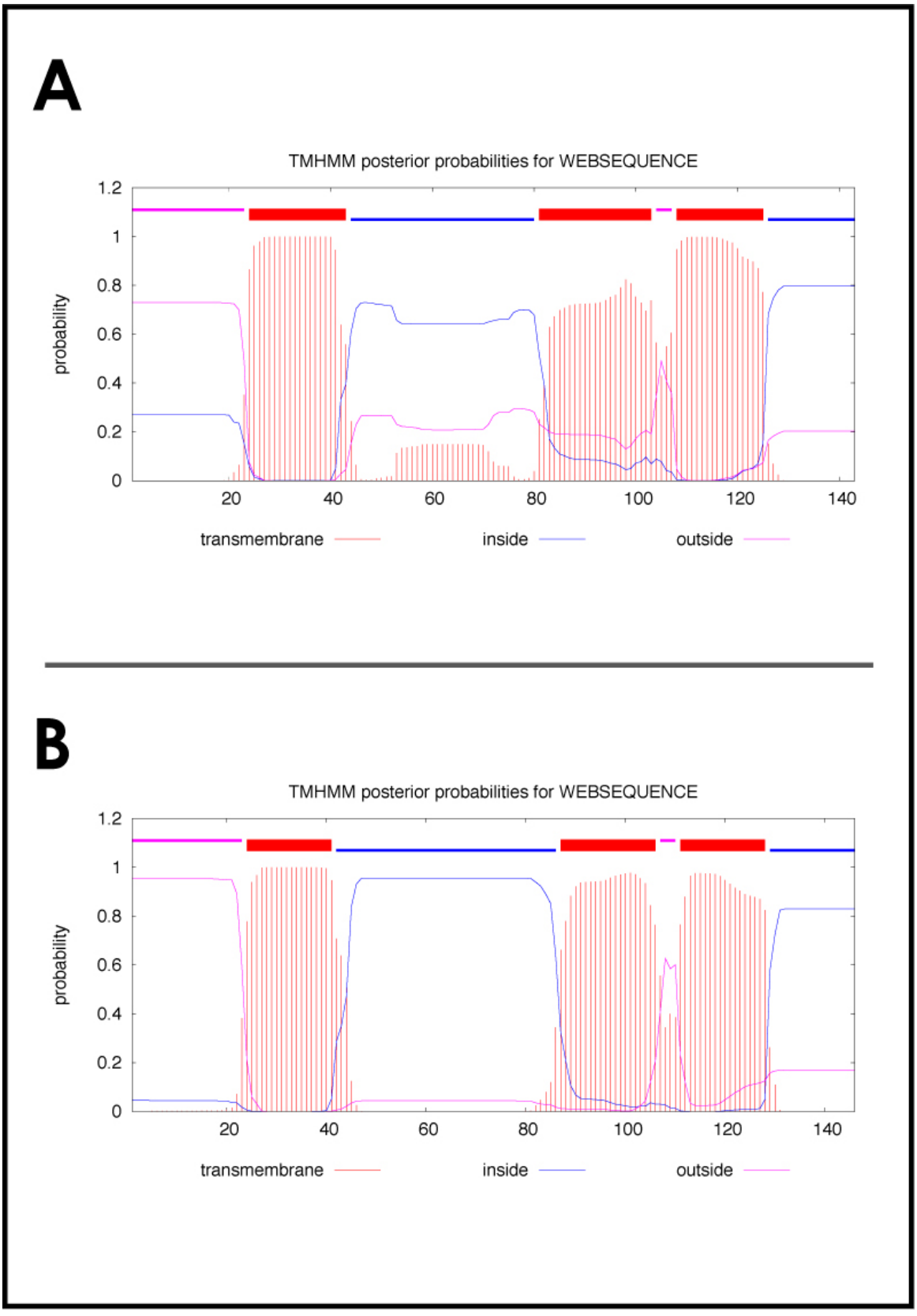


Figure A2(3). Transmembrane domain predictions for ScCOP (**A**) and AtCOP (**B**) generated using TMHMM. The total number of TMDs predicted is three.

VIII APPENDIX 3

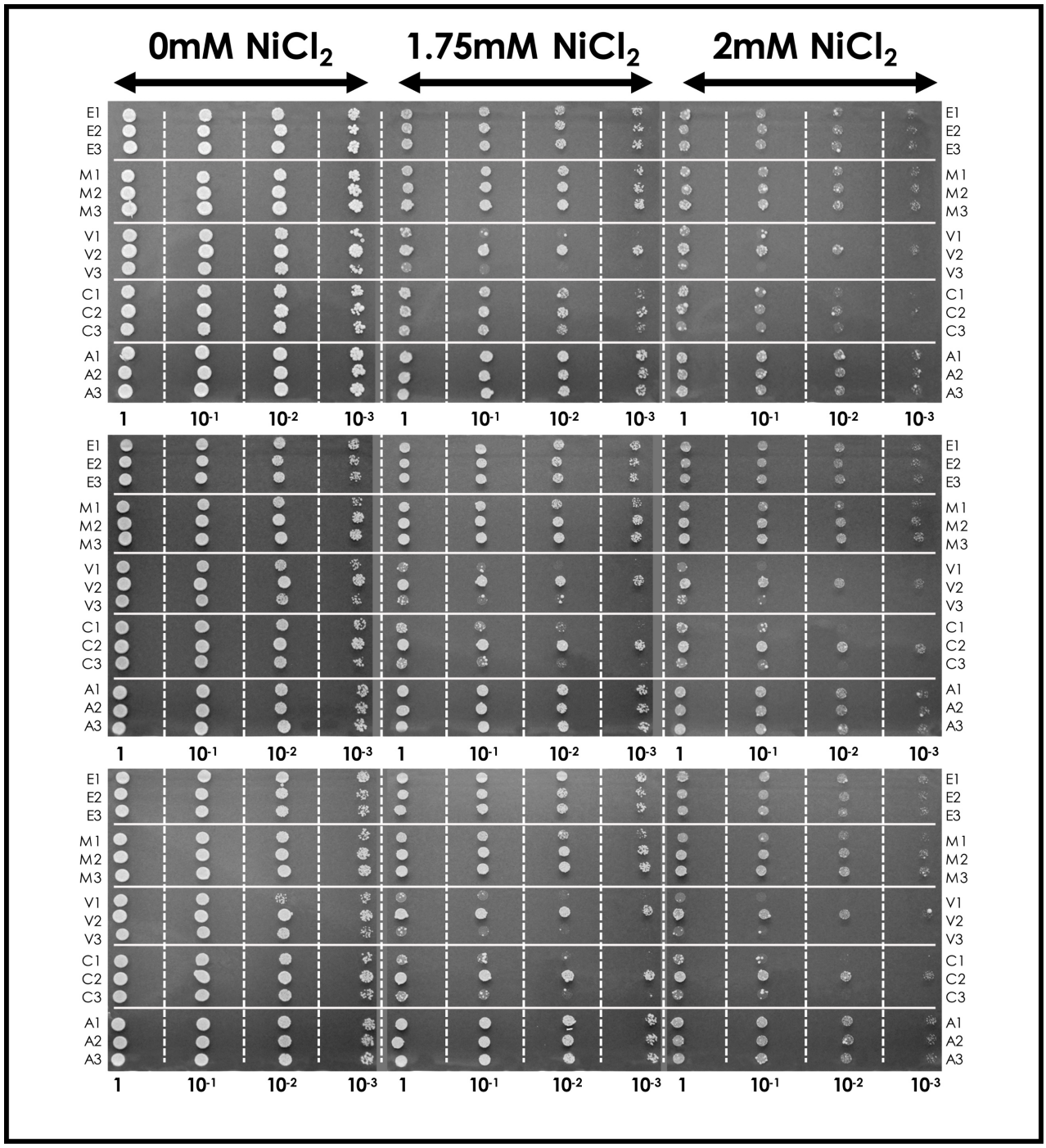


Figure A3. Nickel sensitivity assays in which confirmed positive *S. cerevisiae* (*zrt1zrt2*) transformants containing the empty pDR195 vector (E), pDR195-*ScMATE* (M), pDR195-*ScVIT* (V), pDR195-*ScCOP* (C) and pDR195-*AtIREG2* (A) were spotted out on three CSM-uracil agar plates containing 0 mM, 1.75 mM or 2.0 mM nickel chloride. Three independent transformants were used for each yeast mutant strain and 3 µL of each culture was plated out. The experiment was repeated three times with freshly-made cultures.